

```

1 CTGCAGCCAA CTTTGTGAC CATCTCCGCA ATGCCTTGGA CGTCCTGCAT
51 AGAGAGCTTT TCCCTTAGGT GCCCAGAGTC CTGGTCAACC TCGTGGACTT
101 CCTGAACCCC ACTATCATGC GCAGGCCAGA GCAGCATGCG CGAGCTGGTG GGGTCAGGCC
151 GCCCAGTGCA GCAGGCCAGA GCAGCATGCG CGAGCTGGTG GGGTCAGGCC
201 GCTATGACAC GCAGGAGGAC TTCTCTGTGG TGCTGCAGCC CTTCTTCCAG
251 AACATCCAGC TCCCTGTCCT GCGCCTTGAA CCACTTGGA GCAAAACAGA
301 GACCCTGGAC CTGAGAGCAG AGATGCCCAT CACCTGTCCC ACTCAGAATG
351 AGCCCTTCCT GAGAACCCCT CGGAATAGTA ACTACACGTA CCCCATCAAG
401 CCAGCCATTG AGAACTGGGG CAGTGACTTC CTGTGTACAG AGTGAAGGC
451 TTCCAATAGT GTTCCAACCT CTGTCCACCA GCTCCGACCA GCAGACATCA
501 AAGTGGTGGC CGCCCTGGGT GACTCTCTGA CTACAGCAGT GGGAGCTCGA
551 CCAAACAACCT CCAGTGACCT ACCCACATCT TGGAGGGGAC TCTCTTGAG
601 CATTGGAGGG GATGGGAACT TGGAGACTCA CACCACACTG CCAACATTTC
651 TGAAGAAGTT CAACCTTAC CTCCTTGGCT TCTCTACCAG CACCTGGGAG
701 GGGACAGCAG GACTAAATGT GGCAGCGGAA GGGGCCAGAG CTAGGGACAT
751 GCCAGCCCAG GCCTGGGACC TGGTAGAGCG AATGAAAAAC AGCCCCGACA
801 TCAACCTGGA GAAAGACTGG AAGCTGGTCA CACTCTTCAT TGGGGTCAAC
851 GACTGTGTGC ATTACTGTGA GAATCCGGAG GCCCACTTGG CCACGGAATA
901 TGTTAGCAC ATCCAACAGG CCCTGGACAT CCTCTCTGAG GAGCTCCCAA
951 GGGCTTTTCG CAACGTGGTG GAGGTCATGG AGCTGGCTAG CCTGTACCAG
1001 GGCCAAGGCG GGAAATGTGC CATGCTGGCA GCTCAGAAAC ACTGCACTTG
1051 CCTCAGACAC TCGCAAAGCT CCCTGGAGAA GCAAGAACTG AAGAAAGTGA
1101 ACTGGAACCT CCAGCATGGC ATCTCCAGTT TCTCCTACTG GCACCAATAC
1151 ACACAGCGTG AGGACTTTGC GGTGTGGTG CAGCCTTTCT TCCAAAACAC
1201 ACTACCCCA CTGAACGAGA GAGGGGACAC TGACCTCACC TTCTTCTCCG
1251 AGGACTGTTT TCACTTCTCA GACCGCGGGC ATGCCGAGAT GGCCATCGCA
1301 CTCTGGAACA ACATGCTGGA ACCAGTGGGC CGCAAGACTA CCTCCAACAA
1351 CTTCACCCAC AGCCGAGCCA AACTCAAGTG CCCCTCTCCT GAGAGCCCTT
1401 ACCTCTACAC CCTGCGGAAC AGCCGATTGC TCCCAGACCA GGCTGAAGAA
1451 GCCCCGAGG TGCTCTACTG GGCTGTCCCA GTGGCAGCGG GAGTCGGCCT
1501 TGTGTGGGC ATCATCGGGA CAGTGGTCTG GAGGTGCAGG AGAGGTGGCC
1551 GGAGGGAAGA TCCTCCAATG AGCCTGCGCA CTGTGGCCCT CTAGGCCCGG
1601 GGGTGGGTCC TCACCTAAA CTCCCTATAG CCACTCTCTT CACCGCCTC
1651 TGCCCCAGCC ACTCCCGGCC ACCAGGACAT GCTTCAATGC CTGGTGCCAT
1701 AGGAAGCCCA GGGGACAGTC ACAACTTCTT GGGGCCTGGG CTTCTTCCAG
1751 GCCTATGCTC CTGGAATGGA TACATTTAAA TAAAGTCCAA AGCTATTTTA
1801 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAA

```

#### FEATURES:

5'UTR: 1 - 175  
Start Codon: 176  
Stop Codon: 1592  
3'UTR: 1595

#### Homologous proteins:

##### Top BLAST Hits

	Score	E
CRA 18000004885276 /altid=gi 464376 /def=sp Q05017 PHLX_RABIT P...	750	0.0
CRA 18000005150386 /altid=gi 3172337 /def=gb AAC40129.1  (AF045...	682	0.0
CRA 18000005121266 /altid=gi 2696236 /def=dbj BAA23813.1  (D636...	665	0.0
CRA 18000005181876 /altid=gi 7498717 /def=pir  T20655 hypotheti...	228	9e-59
CRA 87000001028586 /altid=gi 7332170 /def=gb AAF60857.1  (AC024...	210	3e-53
CRA 18000005040393 /altid=gi 7508802 /def=pir  T26083 hypotheti...	205	7e-52
CRA 89000000196200 /altid=gi 7293699 /def=gb AAF49069.1  (AE003...	200	3e-50
CRA 89000000199135 /altid=gi 7297015 /def=gb AAF52285.1  (AE003...	189	6e-47
CRA 18000004979533 /altid=gi 7499049 /def=pir  T16060 hypotheti...	161	2e-38
CRA 18000005184633 /altid=gi 7506410 /def=pir  T24016 hypotheti...	152	9e-36
CRA 18000005184632 /altid=gi 7506411 /def=pir  T24015 hypotheti...	122	1e-26
CRA 18000005182912 /altid=gi 7500588 /def=pir  T21835 hypotheti...	119	6e-26
CRA 87000001028649 /altid=gi 7332235 /def=gb AAF60922.1  (AC006...	111	2e-23

FIGURE 1, page 1 of 2

BLAST dbEST hits:

gi 2079883 /dataset=dbest /taxon=9606 ...	724	0.0
gi 11593761 /dataset=dbest /taxon=960...	670	0.0
gi 7037501 /dataset=dbest /taxon=9606...	654	0.0
gi 12241943 /dataset=dbest /taxon=96...	632	e-179
gi 10367787 /dataset=dbest /taxon=960...	575	e-161
gi 9969781 /dataset=dbest /taxon=960...	547	e-153
gi 7667765 /dataset=dbest /taxon=9606...	531	e-148
gi 12241345 /dataset=dbest /taxon=96...	519	e-145
gi 2080047 /dataset=dbest /taxon=9606 ...	468	e-129
gi 7946640 /dataset=dbest /taxon=960...	323	7e-86

**EXPRESSION INFORMATION FOR MODULATORY USE:**

library source:

Expression information from BLAST dbEST hits:

gi 2079883	Mixed (melaonocyte, fetal heart, pregnant uterus)
gi 11593761	Kidney
gi 7037501	Whole blood
gi 12241943	Lung, normal
gi 10367787	brain glioblastoma
gi 9969781	Prostate
gi 7667765	Colon
gi 12241345	Lung, normal
gi 2080047	Mixed (melaonocyte, fetal heart, pregnant uterus)
gi 7946640	Colon

Expression information from PCR-based tissue screening panels:

Leukocyte

```

1 MRELVGSGRY DTQEDFSVVL QPFFQNIQLP VLALEPLGSK TETLDLRAEM
51 PITCPTQNEP FLRTPRNSNY TYPIKPAIEN WGSDFLCTEW KASNSVPTSV
101 HQLRPADIKV VAALGDSLTT AVGARPNSS DLPTSWRGLS WSIGGDGNLE
151 THTTLPNLIK KFNPYLLGFS TSTWEGTAGL NVAAEGARAR DMPAQAWDLV
201 ERMKNSPDIN LEKDWKLVTL FIGVNDLCHY CENPEAHLAT EYVQHIQQAL
251 DILSEELPRA FVNVVEVMEL ASLYQGQGGK CAMLAAQNNC TCLRHSQSSL
301 EKQELKKVNW NLQHGISSFS YWHQYTQRED FAVVVPFFQ NTLTPLNERG
351 DTDLTFFSED CFHFSDRGHA EMAIALWNNM LEPVGRKTS NNFTHSRAKL
401 KCPSPEPYL YTLRNSRLLP DQAEAEPEVL YWAVPVAAGV GLVVGIIGTV
451 VWRCRRGRR EDPPMSLRTV AL

```

#### FEATURES:

##### Functional domains and key regions:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

Number of matches: 5

```

1      69-72 NYTY
2     127-130 NNSS
3     128-131 NSSD
4     289-292 NCTC
5     392-395 NFTH

```

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
cAMP- and cGMP-dependent protein kinase phosphorylation site

386-389 RKTT

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 7

```

1      7-9 SGR
2     64-66 TPR
3     135-137 SWR
4     326-328 TQR
5     365-367 SDR
6     412-414 TLR
7     466-468 SLR

```

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 6

```

1     12-15 TQED
2     39-42 SKTE
3     56-59 TQNE
4    172-175 STWE
5    298-301 SSLE
6    326-329 TQRE

```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 8

1	123-128	GARPNN
2	138-143	GLSWSI
3	144-149	GGDGNL
4	147-152	GNLETH
5	179-184	GLNVAA
6	278-283	GGKCAM
7	441-446	GLVVGI
8	445-450	GIIGTV

[6] PDOC00009 PS00009 AMIDATION  
Amidation site

Number of matches: 2

1	384-387	VGRK
2	457-460	GGRR

[7] PDOC00016 PS00016 RGD  
Cell attachment sequence

349-351 RGD

[8] PDOC00200 PS00228 TUBULIN\_B\_AUTOREG  
Tubulin-beta mRNA autoregulation signal

1-4 MREL

**Membrane spanning structure and domains:**

Helix	Begin	End	Score	Certainty
1	164	184	0.694	Putative
2	432	452	1.956	Certain

# **BLAST Alignment to Top Hit:**

>CRA|18000004885276 /altid=gi|464376 /def=sp|Q05017|PHLX\_RABIT  
 PHOSPHOLIPASE ADRAB-B PRECURSOR /dataset=nraa  
 /length=1458  
 Length = 1458

Score = 750 bits (1915), Expect = 0.0  
 Identities = 368/502 (73%), Positives = 407/502 (80%), Gaps = 33/502 (6%)

Query: 1 MRELVGSGRYDTQEDFSVVLQPPFQNIQLPVLA----- 33  
 +RELV SGRYDT+EDFSVVLQPPF +IQLPVL  
 Sbjct: 955 LRELVESGRYDTREDFSVVLQPPFHSLQPLVLDGRDLTSFFAPDCVHPNQKFHSQLSRA 1014

Query: 34 -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTPIKPAIENWGSDFLCT 88  
 LEPLG KT+ LDL A + +TCPTQNEPFLRT RNS+YTYP +PA+ENWGSDFLCT  
 Sbjct: 1015 LWRNMLEPLGGKTDALDLTAAILTLCPTQNEPFLRTFRNSDYTYPSRPAVENWGSDFLCT 1074

Query: 89 EWKASNSVPTS SVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSWSIGGDGN 148  
 W AS VP SVH+L+P DIKVAALGDSLTT A+GARP+NSSD P WRGLSWSIGGDG  
 Sbjct: 1075 AWNASRGVPNSVHELQPGDIKVAALGDSLTTAMGARPNSSDPPMFWRGLSWSIGGDGA 1134

Query: 149 LETHTTLPNLLKKNPYLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVERMKNNSPD 208  
 LETHTTLPNLLKKNPY +LGFST T EGT GLNVA +GARA+DMPAQA DLVERM+NSP+  
 Sbjct: 1135 LETHTTLPNLLKKNPYLLGFSTGTLEGTMLNVAVQGARAQDMPAQARDLVERMRNSPE 1194

Query: 209 INLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVEVM 268  
 I+LEKDWKLVTLF+G NDLCH+CENPE EYVQHIQQALD+L EELPR FVNVVEVM  
 Sbjct: 1195 IDLEKDWKLVTLFVGNDLCHFCENPEGSSEGEYVQHIQQALDVLVEELPRTFVNVVEVM 1254

Query: 269 ELASLYQGGGKCA-MLAAQNNCTCLRHSQSSLEKQELKKVNNLQHGSISSFSYWHQYTQ 327  
 ELA L+Q QGG+CA +LAAQ++CTC ++SQSS+E QELKKVNNLQ G+S SY HQY Q  
 Sbjct: 1255 ELAGLHQDQGGRCATLLAAQSHCTCFKYSQSSVEMQELKKVNNLQSGLSRLSYSHQYVQ 1314

Query: 328 REDFAVVVQPPFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNMLEPVGRK 387  
 REDFAVVVQPPFQNTL PLN RGDIDLTTFFS+DCFHFS+RGHAEMAIALWNNMLEPVG K  
 Sbjct: 1315 REDFAVVVQPPFQNTLVPLNGRGDTDLTFFSDDCFHFSERGRHAEMAIALWNNMLEPVGHK 1374

Query: 388 TTSNNFTHSRAKLKCPSPSPYLYTLRNSRLLPDQAEAEPEVLYWAVPVAAGVLVVGII 447  
 TTSNNFT+SR KLKCPSP+SPYLYTLRNSRLLPDQAE P VLYWAVPVAAG GL++GI+  
 Sbjct: 1375 TTSNNFTYSRTKLKCPSPSPYLYTLRNSRLLPDQAEADPTVLYWAVPVAAGAGLLIGIL 1434

Query: 448 GTVVWRRCRRGGRREDPPMSLRT 469  
 V R R REDPP+SL T  
 Sbjct: 1435 AMVAGRGMRCRPREDPPLSLST 1456

Score = 334 bits (847), Expect = 2e-90  
 Identities = 191/445 (42%), Positives = 251/445 (55%), Gaps = 33/445 (7%)

Query: 4 LVGSGRYDTQEDFSVVLQPPFQNIQLPVLA-----LEPLGSKTETL 44  
 L+ S +Y+TQE F+VV QPPF L L +EP+G K E  
 Sbjct: 264 LLASSKYNTQESFAVVFQPPFYESSLSALLAEPPLQDPTTLALSLWNRMMPIGRKEEPF 323

Query: 45 DLRAEMPITCPTQNEPFLRTPRNSNY----TPIKPAIENWGSDFLCTEWKASNSVPTSV 100  
 + P+ CPTQ P+L T RNS + P G++ C + S+SVPTSV  
 Sbjct: 324 SEKERKPLRCPTQESPYLFTYRNSGQLTRVSQPQKGLEVREGTEIRCPDKDPSDSVPTSV 383

Query: 101 HQLRPADIKVVAALGDSLTTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLEHTTLPNI 158  
 H+L+PADIKV+ A+GDSLTT GA +P N D+ T +RGLSWS+GGD N+ T TTL NI  
 Sbjct: 384 HRLKPADIKVIGAMGDSLTTAGNGAGSQPGNILDVLTQYRGLSWSVGGDQNIISTVTTLANI 443

Query: 159 LKKFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKNNSPDINLEKDW 215  
 L++FNP L GFS T T A N A GARA + QA LV MKN IN ++DW

FIGURE 2, page 3 of 7

Sbjct: 444 LREFNPSLQGFVSGTGRETTSQAFFNQAVAGARADGLIPQAQRLVALMKNDTRINFQEDW 503

Query: 216 KLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVVEVME---LAS 272  
K++T+FIG NDLC +C +P + + +I ALDIL E+PRAFVN+V+V+E L

Sbjct: 504 KIITVFIGGNDLDCFCNDPVRYSQNFDTNIGTALDILHAEIPRAFVNLVKVLEISKRE 563

Query: 273 LYQGQGGKCAMLAAQNNTCT-LRHSQSSLEKQELKKVNWNLQHGSSFSYWHQYTQREDF 331  
LYQ C + ++ C C L+ +S E L + Q +Y R+DF

Sbjct: 564 LYQETKVSCPRMILRSLCPCVLKFDNDSTEIASLIETIKEYQERTQQLIDSGRYDTRDDF 623

Query: 332 AVVVQPPFQNTLTPLNERGDTDLTFFSEDCFHFSRGAEMAIALWNNMLEPVGKTTSN 391  
VV+QPPF+ P + G D +FF+ DCFHFS + HA A ALWNNMLEPVG+KTT N

Sbjct: 624 TVVLQPPFEKVNMPKTQDGLPDNSFFAPDCFHFSKAAHAHAASALWNNMLEPVGQKTTN 683

Query: 392 NFTHSRAKLKCPSPESPYLYTLRNS 416  
+F + CP+ P+L T +NS

Sbjct: 684 DF-EGAVNITCPNQVWPFLSTYKNS 707

Score = 323 bits (819), Expect = 3e-87  
Identities = 181/456 (39%), Positives = 261/456 (56%), Gaps = 51/456 (11%)

Query: 2 RELVSGGRYDTQEDFSVVLQPPFQNIQLPVL----- 33  
++L+ SGRYDT++DF+VVLQPPF+ + +P

Sbjct: 609 QQLIDSGRYDTRDDFTVVLQPPFEKVNMPKTQDGLPDNSFFAPDCFHFSKAAHAHAASAL 668

Query: 34 ----LEPLGSKTETDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCTE 89  
LEP+G KT D + ITC P Q PFL T +NS ++ +G+ C +

Sbjct: 669 WNNMLEPVGQKTTNDFEGAVNITCPNQVWPFLSTYKNS-----VQFGTWLPCRD 719

Query: 90 WKASNSVPTSVHQLRPADIKVVAALGDSLTTAVG--ARPNNSSDLPTSWRGLSWSIGGDG 147  
S S PTSVH LRPADI+VVAALGDSL T +G ++PN+ SD T +RGLS+S GGDG

Sbjct: 720 RSPSASPPTSVHALRPADIQVVAALGDSLTAGIGIGSKPNDLSDGTTQYRGLSYSSGGDG 779

Query: 148 NLEHTTLPNILKKFNPYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMK 204  
+L+ TTLNLPN++FN L+GF+ T + + A N A GA+ARD+ +Q LV+RMK

Sbjct: 780 SLDNVTTLPNILRQFNSNLMGFVAVGTGDASGTNAFFNQAVPGAKARDLMSQVQTLVQRMK 839

Query: 205 NSPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNV 264  
+ +N ++DWK++T+ IG +DLC YC + + A + H++ ALD L E+PRA VN+

Sbjct: 840 DDHRVNFQEDWKVITVQIGASDLCYCTDSNLYSAANFYDHLRDALDHLREVPRALVNL 899

Query: 265 VEVME---LASLYQGQGGKCAMLAAQNNTCT-LRHSQSSLEKQELKKVNWNLQHGSSFS 320  
V+ M ++ G KC + A C C L ++S E L+ + Q +

Sbjct: 900 VDFMNPVTRQVFLGNPDKCPVQASALCNCVLSPRENSYELARLEALAQAYQSSLREL 959

Query: 321 YWHQYTQREDFAVVVQPPFQNTLTPLNERGDTDLTFFSEDCFHFSRGAEMAIALWNNM 380  
+Y REDF+VV+QPPF + P+ + G D +FF+ DC H + + H++++ ALW NM

Sbjct: 960 ESGRYDTREDFSVVLQPPFHSLQLPVLQDGRDLTSFFAPDCVHPNQKFHSQLSRALWRNM 1019

Query: 381 LEPVGKTTSNFTHSRAKLKCPSPESPYLYTLRNS 416  
LEP+G KT + + T + L CP+ P+L T RNS

Sbjct: 1020 LEPLGGKTDALDLT-AAITLTCPTQNEPFLRTRNS 1054

Score = 137 bits (341), Expect = 3e-31  
Identities = 107/338 (31%), Positives = 161/338 (46%), Gaps = 42/338 (12%)

Query: 85 FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNSSDLPTSWRGLSWSI 143  
F C + SVP+ SVH LRP+DIK VAA+G+ T + T R +

Sbjct: 46 FPCDPKTLAESVPSESVHSLRPSDIKFVAAIGNVETAPDSGADDLEEQDGTEKRPEQACM 105

Query: 144 GGDGNLEHTTLPNILKKFNPY-LLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVER 202

FIGURE 2, page 4 of 7

```

G             T L I+ + +F+P L+          +   G A D+  QA +LV
Sbjct: 106 G-----VVTVLSDIIGRFSPSALMPLCPET-----RLVPRGG-AEDLWMQATELVRS 151

Query: 203 MKNSPDINLEKDWKLVTLFIGVNDLCHYCENPEAH-LATEYVQHIQQALDILSEELPRAF 261
M+ +P ++ E DWKL+ +F          C C + +   L   + + + LD L +E+P+AF
Sbjct: 152 MRENQQLDFEHDWKLINVFFSNTSQCFPCPSAQQKGLVLGGMDKLTRTLDYLQQEVPKAF 211

Query: 262 VNVVEVMELASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKV--NWNLQHGSISS 319
VN+V++ ELA+   + + G   + A   C CLR          E +L KV   W+          S
Sbjct: 212 VNLVDLSELAAFSRWRQG-AQLSPAAPCRCLR-----ETSQTLTKVLTQWSYLEAWDSL 264

Query: 320 SYWHQYTQREDFAVVVQFFF-QNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWN 378
+Y +E FAVV QFFF +++L+ L                               +   +A++LWN
Sbjct: 265 LASSKYNTQESFAVVFQPFYESSLSALLAEPPL-----QDPTTLALSILWN 310

Query: 379 NMLEPVGRKTTSNFTHSRAKLKCPSPESPYLYTLRNS 416
M+EP+GRK   +   R L+CP+ ESPYL+T RNS
Sbjct: 311 RMMEPIGRKEEPPS-EKERKPLRCPTQESPYLFTYRNS 347

>CRA|18000005150386 /altid=gi|3172337 /def=gb|AAC40129.1| (AF045454)
      phospholipase B [Cavia porcellus] /org=Cavia porcellus
      /taxon=10141 /dataset=nraa /length=1463
      Length = 1463

      Score = 682 bits (1741), Expect = 0.0
      Identities = 348/505 (68%), Positives = 389/505 (76%), Gaps = 38/505 (7%)

Query: 1      MRELVGSGRYDTQEDFSVVLQPFQNIQLPVLA----- 33
MRELV SGRYDT+EDFSVVLQPF NI+LP+L
Sbjct: 954 MRELVESGRYDTREDFSVVLQPFFLNIRLPILEDGRPDTSFFAPDCINPGQKFHSQLSRA 1013

Query: 34      -----LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTIPIKPAIENWGSDFLCT 88
LEP+GSKT+TLDL A++ + CPTQ EPFLRTP+NS+YYP KPAIENWGSDFLCT
Sbjct: 1014 LWVNMLEPVGSKTDTDLTADISLPCPTQEEPFLRTPQNSDYTYPTKPAIENWGSDFLCT 1073

Query: 89      EWKASNSVPTS SVHQLRPADIKVVAALGDSLTTAVGARPNSSDLPTS WRGLSWSIGGDGN 148
EWK SNSVPTS SVH+L+PADIKVVAALGDSLTTAVGAR +NSSDL SWRGLSWSIGGDG
Sbjct: 1074 EWKPSNSVPTS SVHKLQPADIKVVAALGDSLTTAVGARASNSSDLLMSWRGLSWSIGGDGA 1133

Query: 149     LETHTTLPNILKKFNPYLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLVERMKNSPD 208
LEHTTTLPNILKKFNP + GFST T E TAG NVA E ARARDMPAQ DLVERMK S +
Sbjct: 1134 LETHTTLPNILKKFNPSIFGFSTGTLEETAGFNVAVEEARARDMPAQARDLVERMKASTE 1193

Query: 209     INLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFFNVVE-V 267
INLE DWKL+TLFIG NDLCHYC+NPE H A EYVQHI+QALDIL EELPRAF+NVV+ +
Sbjct: 1194 INLEMDWKLLITLFIGSNDLCHYCDNPENHSAEYVQHIRQALDILYEELPRAFINVVDII 1253

Query: 268     MELASLYQGQGGKC-AMLAAQNNCTCLRHSQSSLEKQELKKVWNWNLQHGSISSFSYWHQYT 326
MELA L+QGQGG C A+L AQ+ C+CLRH SS QELKKV WNLQ +S SY +YT
Sbjct: 1254 MELAGLHQGGGHCTALLPAQSTCCLRHFPSSPVIQELKKVTWNLQSDMSRLSYQEKYT 1313

Query: 327     QREDFAVVVQPFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNMLEPVGR 386
QREDFAVVVQPFQNTL PL++ G TD TFFSEDC HFS+RGHAEMAIALWNNMLEPVG
Sbjct: 1314 QREDFAVVVQPFQNTLIPLDKLGSTDPTFFSEDCLFHFSERGRHAEMAIALWNNMLEPVGH 1373

Query: 387     KTTSNFTHSRAKLKCPSPESPYLYTLRNSRLLPDQAEAEVLYWAVPVAAG----VGL 442
KTT NNF++R KLKCPSP ESPYLYTL+NS LP Q E+A V V AA VGL
Sbjct: 1374 KTTFNFTYNTKLCPSSTESPYLYTLQNSLSLPVQTEKASGVAPGIVSAAAAGLLVGL 1433

Query: 443     VVGIIIGTVVWRCRRGGRREDPPMSL 467
+VGI+ +W R +++ PP S+
Sbjct: 1434 IVGILAVSLWSSFRRRQKKSPPEV 1458

```

FIGURE 2, page 5 of 7

Score = 348 bits (884), Expect = 7e-95  
Identities = 199/442 (45%), Positives = 257/442 (58%), Gaps = 31/442 (7%)

Query: 4 LVGSGRYDTQEDFSVVLQPPFQNIQLPV-----LAL-----EPLGSKTETLDL 46  
L+ S ++ QE F+VV QPFF + PV LAL +P+G K E  
Sbjct: 265 LLASSENQESFAVVFQPPFFYEVSSPVEEPSQDPTTLALSLLWNNMMKPVGQKDEPFST 324

Query: 47 RAEMPITCPTQNEPFLRTPRNSNYTYPI----KPAIENWGSDFLCTEWKASNSVPTSVMH 102  
P+ CP+Q P+L T RNSNY + + E G++ C + S+S PTSVMH+  
Sbjct: 325 IERRPMKCPSESYPFLTYRNSNYQSRLLRQQRQHKEREGTEIRCPDKDPSDSTPTSVHR 384

Query: 103 LRPADIKVVAALGDSLTAVGA--RPNNSSDLPTSWRGLSWSIGGDGNLEHTTLPNLIK 160  
L+PADIKV+ ALGDSLT GA RP N D+ T +RGLSWSIG D N+ + TTLPNLIK+  
Sbjct: 385 LKPADIKVIGALGDSLTAGNGAGSRPGNILDVLTEYRGLSWSIGADHNISVTTLPNLIK 444

Query: 161 KFNPHYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKNSPDINLEKDWKL 217  
+FNP L GFST T + A N A GARA D+ QA LV+ MKN IN E+DWK+  
Sbjct: 445 EFNPSLKGSTGTGKANSVGAFNQAVAGARAGDLIPQARTLVDLMKNHSTINFEEDWKI 504

Query: 218 VTFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVNVEVME---LASLY 274  
+T+FIG NDLC +C +P + + +I+QALDIL E+PRAFVN+V+V++ L LY  
Sbjct: 505 ITVFIGGNDLCDFCSDPVTNSPENFTDNIRQALDILHAEPVRAFVNVMKVQLQIVNLRELY 564

Query: 275 QGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKNWNLQHGISSFSYWHQYTQREDFAV 333  
+ C L +N C C L +S E + L +N Q +Y REDF V  
Sbjct: 565 KDSRVSCPRLILRLNLCRCVLLPDDNSTELESIDINKKYQERTHQLIESGRYDTREDFTV 624

Query: 334 VVQPPFQNTLTPLNERGDTDLTFFSEDCFHFSRGAEMAIALWNNMLEPVGRKTTNNF 393  
V+QPPF+ P G D T F+ DCFHFS + HA A ALW NMLEPVG+KTT NNF  
Sbjct: 625 VLQPPFEKVDIPKTSEGLPDNTSFAPDCFHFSKTHARAASALWKNMLEPVGQKTTQNNF 684

Query: 394 THSRALKCPSPESPYLYTLRN 415  
+S + CP+ PYL T +N  
Sbjct: 685 ENS-IDIICPNQAFPYLSTYKN 705

Score = 314 bits (795), Expect = 2e-84  
Identities = 178/455 (39%), Positives = 259/455 (56%), Gaps = 51/455 (11%)

Query: 3 ELVSGSGRYDTQEDFSVVLQPPFQNIQLPVLA----- 33  
+L+ SGRYDT+EDF+VVLQPPF+ + +P +  
Sbjct: 609 QLIESGRYDTREDFTVVLQPPFEKVDIPKTSEGLPDNTSFAPDCFHFSKTHARAASALW 668

Query: 34 ---LEPLGSKTETLDLRAEMPITCPTQNEPFLRTPRNSNYTYPIKPAIENWGSDFLCTEW 90  
LEP+G KT + + I CP Q P+L T +N IE G+ C E  
Sbjct: 669 KNMLEPVGQKTTQNNFENSIDIICPNQAFPYLSTYKNG-----IEGHGTWLTCTER 719

Query: 91 KASNSVPTSVMHQLRPADIKVVAALGDSLT--TAVGARPNSSDLPTSWRGLSWSIGGDGN 148  
S S PTSVMH LRPAD++VVAALGDSLT + +G++P + +D+ T +RGLS+S GGDG+  
Sbjct: 720 TPSASPTSVHALRPADVRVVAALGDSLTAGSGIGSKPGDLADVITQYRGLSYSSGGDGS 779

Query: 149 LETHTTLPNLIKKNFNPHYLLGFSTSTWEGT---AGLNVAEAGARARDMPAQAWDLVERMKN 205  
L TTLPNLIK++FN L G++ T + + A LN A GA+A ++ +Q LV++MK+  
Sbjct: 780 LMNVTTLPNLIKREFNSNLTYAVGTGDASNTNAFLNQAVPGAKAEELMSQVKTLVQKMKD 839

Query: 206 SPDINLEKDWKLVTFIGVNDLCHYCENPEAHLATEYVQHIQQALDILSEELPRAFVN 265  
P IN +DWK++T+ IG NDLC++C + + + + + H+ ALDIL E+PRA VN+V

FIGURE 2, page 6 of 7



Sbjct: 840 DPRINFHEDWKVITVLIGTNDLCNHCTDLDLYSSANFFNHLNLDILHREVPRALVNLV 899

Query: 266 EVME---LASLYQGQGGKCAMLAAQNNCTC-LRHSQSSLEKQELKKVNWNLQHGISSFSY 321  
+ M + ++ G KC + A C C L ++S E + + Q +

Sbjct: 900 DFMNPSIMRQVFLGNPDKCPVQQASILCNCVLSLRENSYELARMDALTRAYQSSMRELVE 959

Query: 322 WHQYTQREDFAVVVQPFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALWNNML 381  
+Y REDF+VV+QPF N P+ E G D +FF+ DC + + H++++ ALW NML

Sbjct: 960 SGRYDTREDFSVVLQPFFLNIRLPILEDGRPDTSFFAPDCINPGQKFHSQLSRALWVNML 1019

Query: 382 EPVGRKTTSNFTHSRAKLKCPSPESPYLYTLRNS 416  
EPVG KT + + T + L CP+ E P+L T +NS

Sbjct: 1020 EPVGSKTDTLDT-ADISLPCPTQEEFPLRTPQNS 1053

Score = 155 bits (389), Expect = 8e-37  
Identities = 110/351 (31%), Positives = 166/351 (46%), Gaps = 48/351 (13%)

Query: 85 FLCTEWKASNSVPT-SVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTS---WRGLS 140  
F C+ K ++P+ SVH L PADIK++AA+GD T N + T WRG

Sbjct: 45 FSCSPKKLGLNMPSESVHTLTLPADIKLIAAIGDMETPPDSGAVNLDTSETEKEPWRGCM 104

Query: 141 WSIGGDGNLEHTTTLPLNILKKFNPYLLGFSTSTWEGTAGLNVAEAGARARDMPAQAWDLV 200  
+ T L +I+ FNP +L + W AA ++ QA +LV

Sbjct: 105 GMM-----TVLSDIISHFNPSVLLPTCPPWRS-----AAVRGGVEELRTQAEELV 149

Query: 201 ERMKNSPDINLEKDWKLVTLFIGVNDLCHYCENPEAHLATEYVQHIQQ---ALDILSEEL 257  
+K +P ++ ++DWKL+ +F LC+ C P AH + ++ + L L +E+

Sbjct: 150 SSLKKNPQLDFQQDWKLVNFFSNASLCYLC--PSAHENGPLMSNMDKLAGILHYLHQEV 207

Query: 258 PRAFVNVEVMEASLYQGQGGKCAMLAAQNNCTCLRHSQSSLEKQELKKVNWNLQHGIS 317  
PRAFVN+V++ E+ ++ + G + C C K + + W+ Q

Sbjct: 208 PRAFVNVLVDLFEVVMAPRWHQGTMLSRPSPEACGC----SGETSKLDTVVMQWSYQETWD 263

Query: 318 SFSYWHQYTQREDFAVVVQPFQNTLTPLNERGDTDLTFFSEDCFHFSRGRHAEMAIALW 377  
S + +E FAVV QPFF +P+ E D T +A++LW

Sbjct: 264 SLLASSSFNDQESFAVVFQPFYEVSSPVEEPPSQDPT-----TLALSLW 308

Query: 378 NNMLEPVGRKTTSNFTHSRAKLKCPSPESPYLYTLRN---SRLLPDQAE 424  
NNM++PVG+K + T R +KCPS ESPYL+T RN SRLL Q +

Sbjct: 309 NNMMPVQKDEPFS-TIERRPMKCPSQESPYLFTYRNSNYQSRLKRQRQ 358

#### Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF00657	Lipase/Acylhydrolase with GDSL-like motif	158.4	1.6e-45	1
PF01347	Lipoprotein amino terminal region	1.9	6.1	1
CE00543	CE00543 steroid_receptor_N10	-0.8	1.5	1

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01347	1/1	155	172 ..	636	653 ..	1.9	6.1
PF00657	1/1	110	233 ..	1	146 []	158.4	1.6e-45
CE00543	1/1	230	254 ..	456	480 ..	-0.8	1.5

```

1  ATTCTGCAGC  CAACTTTGTT  GACCATCTCC  GCAATGCCTT  GGACGTCCTG
51  CATAGAGAGG  TGGGTGGGGG  GCTTCCACAA  GCTGGTAACA  GCTCAAGCAT
101  GGTGAGGGTG  AAGGTGGATG  GGGGGAAGA  ATGAGAGAAG  AACCCTTTC
151  TCTCAAGGAG  ACAGCCAAGG  GCATGGANN  NNNNNNNNN  NNNNNNNNN
201  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
251  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
301  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
351  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
401  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
451  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
501  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
551  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
601  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
651  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
701  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
751  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
801  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
851  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN
901  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NNNNNNNNN  NGTGCCATTG
951  CCGCTGCAGC  CCCTTTGGGT  GGCACCCATG  GAGTTGTGTG  ATGTACGGCC
1001  TGTAAAGCCT  TACGAGGTAG  CCCTGTATAG  ACTCCTCCCC  AGAATCAAC
1051  TCCAGAAAGA  CCAAGCTGGA  TTGCTAAAGG  AACCCATTCC  TAGGGGCCCT
1101  GAGACAGCCC  CAGGAAGAAG  TGCCTGGAGC  CCCCCTCTCA  TCTGCAGCTT
1151  TTCCCTTAGG  TGCACAGAGT  CCTGGTCAAC  CTCGTGGACT  TCCTGAACCC
1201  CACTATCATG  CGGCAGGTGT  TCCTGGGAAA  CCCAGACAAG  TGCCAGTGC
1251  AGCAGGCCAG  GTAGGCAGGT  CCTGGCTGTC  CCCACACTGG  AGATGCCCTC
1301  ACCTCCTGGT  CTGGCCACA  TGCAGTGGTG  ATGCCTCAGG  GTCTTTGTGA
1351  CTTGGTCTAT  CCATGTGTCC  AAGTCTGTAA  AGGAGGACTT  CTGCCAGAAC
1401  GTCCCTTCC  AGAGGCTGGA  GCCATGACTC  CCCTGTTACC  CAACTCAAG
1451  GTGCCTGGCA  GGAATTCTA  TGATACCAGG  CAGCCACAGA  GGGGAGGGAT
1501  CAAAGTTGGG  ACAGAGGCTG  GTGTTTGAGA  GACAGGATAG  CCTAGACTGT
1551  GAACATGGGC  AGTGGTTAGG  GATGTAGACA  TATGTGGTCA  AACTGTAACA
1601  GAAAGCAAGG  AAAAGGTACA  AGCAACTCAG  TTACCTTTAG  GGGAAAGAAG
1651  GAATTAGGAG  GGACACAGGG  AGCTTCAAAC  TGGGAGTGTT  TTGTTTCTTA
1701  AACTGGGCCA  TAAGTACATG  GATGTGTGTT  TTATTATCT  TTATATCTTA
1751  CACATCTATT  TACTCAGCAA  ATCTTACAGA  ACTTCTGTG  TACCAGGCAT
1801  TGTTTCAAGT  GCCTTAGCAA  TCTCTCTCT  AAGTAGATGT  GATGGGTGTG
1851  AAATAATTCA  TGATGAAACC  AAAGGGGACA  CAGTAGGGCA  CTCATGTGAA
1901  AGAAGGAGAG  GTCTAAGGCA  TAGCATCAGA  GGCCCCAAA  TATCAGCTCC
1951  AACACCAGAG  GATGCATTTT  CTTTTTAATT  AAACACTAAA  TTTTCACTGC
2001  CCAAATTCAT  TTGCTCAGCT  GAATAATCGG  TTGCAGGCC  AGCACCTGCA
2051  GTCCAACACT  TGTGCTCTGT  TGGTATGAGA  GGGTGCTCAT  TCCCACGCTG
2101  GCTCCCTCCC  TCGGGCCATC  TCCAGTGCCC  CTGCCAGGCC  TGAAGCCTGC
2151  CCCTGAGCAT  GTGCCCCAGA  GCCTCAAGGC  TTGAGTGCTC  CTAAACCAGG
2201  GCGGGAGGGA  GCCTCTCCAC  CCCTCCCCTG  AACCTGGGCA  ATCAGAACCA
2251  GCCCCTGATG  GAAGCCTGAG  CTCTGGGGCC  TCCTGCCTCC  CCCTCTTTGT
2301  GCAGCGTTT  GTGTAAGTGC  GTTCTGACCC  TCGGGGAGAA  CTCCCAAGAG
2351  CTAGCCAGGC  TGGAGGCCTT  CAGCCGAGCC  TACCGGGTAA  GACCAAGAAG
2401  GGCACCATGC  TGTGTCCTCT  CCCCTACGTT  CACTCTAACA  CACAGCCAG
2451  AGCCCTAGA  GGAGGCACAC  AGGGAAGGAA  AAGCTGGTCA  GGGATTGTGG
2501  GGAGACGGGG  AGCAGCCTGG  GTGCCTTCCT  CTGTCTCACG  TGAATGTGGT
2551  GTCTCAGGTG  CCCTGGTTGG  AATCATCCCA  GTAGGATCCA  GGTGGAAAAG
2601  CCCTCATGGC  CCAGCTACCG  TTGAGGGCTT  AACCCCAACT  CCTGGCCCGT
2651  AGCCCTGGAT  GCCTCATGAG  ACCACCTTTC  CCTCCCCAC  TCCCACTCCA
2701  AAGGCAGGTG  CCGAGCCTCT  GGAGGTTCTT  CCCAGGTTTT  TATCCCTTTT
2751  GGGACTTCTT  GCCTAGCCCT  TCAGAGAGAG  TAGTCTACTT  ACAATCAAAA
2801  CAAAAGGTG  ACCCAACTCG  TTTCCAAATT  CTCTGGAAAG  GGAATGCCCC
2851  TCAGGTGATT  TGTGTTCTCA  AGGGAAAGGC  TGAGTCGGCC  CCTCCATCCA
2901  GGGAGATGGA  CTGCCACCA  CCCCTACTCT  TGCTCTACTG  GGTCTGGGCG
2951  CCACCCAGGG  CCTGGGCTGA  AGACCTGTG  CATGTGTCCC  CAGAGCAGCA
3001  TGCGCGAGCT  GGTGGGTGCA  GGCCGCTATG  ACACGCAGGA  GGAATCTCTT
3051  GTGGTGTGTC  AGCCCTTCTT  CCAGAACATC  CAGCTCCCTG  TCCTGGTGGT
3101  ATGTCCCTCG  CCCTCGCCCA  TGGTACTCTT  TTAGAGGAAG  AAATGCAAGG

```

FIGURE 3, page 1 of 33

3151	CAGAATTGCC	AGTTGCTTCC	ACGAGCATGT	GCATAAAATG	GGAAGACAC
3201	AGCTCTCCAG	ACGCTGNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
3751	TTGTGGGATC	GCTCTGATCT	CTGTGTTAAG	TGAATGGGCC	CTGTGGTGGC
3801	TGGTGACCTG	GAGCATCTCA	GGGGAAGGAA	GGTGTGAGT	GGTCAGTCCC
3851	AGGGCCAAGT	CCGCTGGTGG	TGGCTCCCTC	TGAACCAATA	GGATCTTGAG
3901	GGGGTATATT	GGTCTCTTTC	AGGATGGGCT	CCCAGATACG	TCCTTCTTTG
3951	CCCCAGACTG	CATCCACCCA	AATCAGAAAT	TCCACTCCCA	GCTGGCCAGA
4001	GCCCTTTTGA	CCAATATGGT	AAAAAAGTG	GGGTGTTCCT	TGTTCTCTGG
4051	GGTTCTAGTC	TAGGGCAGGG	CACGAGCCCC	TATAGATATG	GTCTTGCAG
4101	GTGAGCTGTA	GGGGGCAGTG	CTGTGTACAT	CTATAAACGT	CTATGCAGTT
4151	GGAAATGCGG	AGTCCTTAAG	AGTCTGCTCA	GCCTGGGCTC	AACTGCACCC
4201	TCTCCTCAGA	GCTTTGAACT	CTGAGGAGGG	ACCTCTCTAC	AGAAATGCAA
4251	GCCCAAACCC	CCATATTTCAT	TCCACTTTCC	CTATGTGCCG	GCCACCATGT
4301	TAGGCAGTTT	AAGCCACGTT	ATCTCATTTA	AGGCTCTGCA	TATCCCTGCT
4351	AGAGAAGCAT	GACGAGCTCT	CAAGGAAGCT	ACTCCCAGAG	AAGCAAGCG
4401	ACTGGCCCAA	AACCCACCG	CTGCGAACTG	GCAGAGCCAG	AAGTGGGAGC
4451	CAAGCCCCCT	GAAATCGAGT	TCTGAGCTTT	CCCCACTGCA	GGATTCTGCC
4501	AGGGAATGTT	CACTTCCATG	GAAACAAACT	ACTACACCCG	TGTCTCTCTT
4551	TTCTTCCCTG	ATCAGCTTGA	ACCACTTGGA	AGCAAAACAG	AGACCCTGGA
4601	CCTGAGAGCA	GAGATGCCCA	TCACCTGTCT	CACTCAGGTA	GTAGGGGAGG
4651	ACCTGCCTGG	CTCCTCTCCA	CAACCAGGCG	CACACAGCTC	GCCCTACCCA
4701	CTTCGCTCTC	CACCAACAGT	TCCCTAGTAC	CCATCTTGCC	CCCTTACTGA
4751	GGCCTGAGAG	ATTTGGAGGA	TGGAGGGGAG	TCCATGAGGA	TGGACAGGGG
4801	AGGTGAGAGG	GGAGACAAGA	GTGCAGCTGT	CATTGGGAAC	AGGAGATGCA
4851	GCAGGGAGAG	GAGGCTGGG	CCCAGCAGA	GGGAGAGGAT	CCCGGTGAGA
4901	AAAGTGGGCT	CCTGAGAGAG	GAAATCAGGA	TGCCAGGAAA	ATGGCAGGAG
4951	GGCTTCTCTT	AGCATGTGGT	TTTGGGCGAG	ATGAAAAAAT	CTGACTGCAG
5001	GTTAGAGGGC	CCAGGCAGGA	GCCAGGCAGG	TTTAAGAGCT	GTGGTTGGAG
5051	AGAGGAGAGC	CTGGATTAGG	GAGATTCCAC	AAGGAAAGGA	TCACAGAGGA
5101	CAGCAGCAAA	GGGCAGAGCC	CAGAGCTGTA	TGGAGGAGGG	ACGAGGGTGG
5151	GCCTACCAGG	ACACGGCAGC	TCCAGGCTCC	TTTTAAGGAG	GAATCCGTAA
5201	GTGGTTGTTA	AGCTTGACTT	CAGGCCTGGG	GTGGGGGCG	GTTCTCATTT
5251	TCTTCAGCTC	CTGTTCTTAG	GCCCGGTCTT	ATGGCTTTTT	AACCAATAAA
5301	GGCCAAGGCC	AGAAAAACCT	CAGCAGCAAT	AAAGCAGAAA	GGCCTGACCC
5351	AATCTGGGAG	GCTGGGTTTC	CCTCCTAGGT	CGGCCACACC	ACCCTCTCCC
5401	ACCCTCCCTG	CTGGGGAATG	GACCTGCAGC	TCCCCATGT	GTCTGCTGGG
5451	AATCCTGAGA	GAGTGGGCAC	CCCTGTTTAC	ATGCTTGCTC	CCTGTCTGCT
5501	GCCTGCCCTA	CCCCAGTCTT	GGGCTCAGGC	TCAGTCTTGT	GTGCCATCAG
5551	CCCCATCAGG	AGAGCAAGAA	TGGCAGGAAG	AAGGCTTGGG	AAGTGAAGAC
5601	AGTCGTAGCA	GAGGGCTCAG	TTGTCTGGTC	TTGTCTTGG	AGCTAAGGAG
5651	ATTGTCAGAT	TCTGCAACAG	CTAGTGCAAC	ACAGATGCCT	CTAGTCCAGG
5701	TGGTCAGGTG	CTGGCCAAAG	GCCTGGAGCA	AAACCTTAGA	GGCCCCTACT
5751	GTGCCAGGTG	TAAACTCTTT	AACTGCTTTC	CTAAGGATGC	CTTGGGGGTT
5801	CTAGGGGAGC	AGCCAGGGAC	CGTGGATAGT	GGGGGCATTT	GGGGACTCAG
5851	AAATAGCCAT	ATTGTAGATA	TTTCAATATT	TTACCAACCT	TATAGCCATA
5901	CTGAATATCA	GCCATGTGAGG	GCCCTTTTCCA	AACTGTCCAC	TCCCTTCCA
5951	TTACATAACA	AAAGCAGCCA	TCATTTGCTC	TTTCTTTCAA	CAAACGTGTA
6001	TTGAGTACTG	AGTTGGAGCC	TAAGCACTGG	GTCAGGGAGA	GCCCTGTCAC
6051	CCTGGGCTTC	GAGGCAACCA	CTTCCAGGCT	TTACCCAGAG	TCAGGCAGAG
6101	ACCCCCAAAA	GGAGGCTGCT	CCACCCAGCA	GCATCTTAAG	CTGAGTGGGC
6151	TCAGTGCCTC	CCTTCTAGAC	AGAGCCCAAT	GGAGCCACTG	CAGTGAATTT
6201	CAGAGGTGAG	CAGATCTCAGC	CTCGTGGAA	CAGTGAAGC	CCAGCCCTGG
6251	TGAAGCTGTT	GCTAAGCAAC	ATTGGAGCCC	ATTCTGAAAG	GGTCCATCTG

FIGURE 3, page 2 of 33

6301	TTGGCCAGCC	CAACTTCACT	GTGTTCTGAG	CATTCTGCAT	TCCTCAGTCC
6351	CATCTGCTCC	CTCCCATGTG	CCTTGGAGTG	ATATAAAAGT	CCACCAGCAT
6401	CTCAGTGTGA	GCTGACAGGG	GCCAGGCAGC	ACCTATTTTT	GTCCTAGATG
6451	TGTCTAAACA	TAGAGGCAAC	AGGCAACAGG	CAAGACGCAG	TGGGGGGCGG
6501	GAGGCAGGAG	GCCGAGATGG	CTGTGAGCAT	GAGCTTTCTC	AGCCTCCTCC
6551	CTTCTCCCAT	CCGCAGTCTA	ACTGCTCATA	CGTTCTGTGT	GCCAGGTAGG
6601	GTGACTTAAC	AGCACGCCAT	GGATTTCTGT	TGTAGTTTCA	AGTTGGACAA
6651	ATTCTTTTAC	AGACAAC'TTT	TGACTAGCCT	TCTGTGGACT	GAGCCTATAC
6701	TCTGCCTTAA	TGGGCTCTCT	GCCCACTCCT	TTCCTAACCC	CAGGGCAGCT
6751	GGCTGAACAC	CTGGTCCTTT	TCTTAGGTTT	CATTCTTTTT	GACCTCTCTG
6801	AAGCCCTTGT	CAAAAGTCAC	CACCTCCCCC	TTGAAATTCA	CTCCTTCCTG
6851	GGTTTGTGGA	CACTAAATCG	CCTTGATTTT	TCTGGTCTTC	TGTTTGCTTG
6901	CCTTTAATGA	CCCTCCTCCT	CCCTTTCCCC	AGTCTTGAAA	ATGTAGATAT
6951	TCTCCAATTT	TCATGTCTCC	ATTCTATTTT	CTTTCCTTTT	TCACTCACTT
7001	TTTGAAACAG	GGTCTTGCTC	CGTCTCCCAG	GCTGGAAGTG	CAGTGGCGCA
7051	ATCACAGCTC	TCTGCAGCTT	TCAACTCCTA	GGCTCAAGCC	ATCCTCCCAC
7101	CTCAGCTTCC	TGAGTAGTTG	GGACTGCAGG	CATGCACCAC	CATGCCCAGC
7151	TAATTTTGGT	TTATTTGTTT	TGGTAGAGGG	GGGGTCTTGC	CATTTTGGCT
7201	CAGGCTGATT	TTGAACCTCG	GNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
7951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8451	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8501	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8551	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8601	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8651	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8701	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8751	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8801	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8851	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8901	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
8951	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9001	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9051	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9101	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9151	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9201	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9251	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9301	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9351	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN
9401	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN	NNNNNNNNNN

FIGURE 3, page 3 of 33

9451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
9501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN  
9551 NNNNNNNNNN NNNNNNNNNN NNNNCAATAA ACCCAGCTAA AAACAAGCCC  
9601 AATAAAACCC AATAAAACCC ATTAGACAGG AACATAGGAG TTGGAAAAAA  
9651 AAGAAAAGAA GGGGAGGGGG AAGAAAGCCC TGAGGCACCC CGGCTGCCTG  
9701 TCTGCCACAA CCCTGGGCTG TAATTGTTCT TGCCATGGCC TCAGTCTGCA  
9751 ACACATTCTA GTGTCTCCTT GACCTCTAGC CCTCTAGCTC TGCCTCCCTT  
9801 TCCCCAACCT GTAGATCTTG TGATCAAATA GATTCAATGA AACACATTGT  
9851 CCAGTTGCAC TCGCAGCACT TCCAAAAAGG TCAAGTTTGT CCTTCCCTCA  
9901 GTGCCTCCCA TTCTGGTCAC GGTAGGACTG ACTCCAGCCC CTGGACCCTA  
9951 AGCTGAGTCT TGGCTCCTTT GACGTGCAGG GAGAATGCCA CTGAGTCTTG  
10001 TCTCTGAGGA CCCTACCTCT CCAAATCTTG CCTCAGTTCC TCAGCAGGTA  
10051 CTACACTGAC TGGCCATGCC ATTCTCTGAT GCTTCACTGC CTCAGCTTCT  
10101 CAAGTCTGTC TCCCCACCTG AGCCAATGTG GAGTTTCTCT CTCTCCTCCT  
10151 CTCATCCTGG CACCTAGAAA TGCTCTCTAA CGCTTGAGCT GCTCAACCAG  
10201 CATGGGTGAC TTGTTTATAG CATGCTCCCA GATCGCCCTC TTTGTTGGTG  
10251 AATGCTCAGG GAATGCTTAC TGTTAACCCG AGACAAGCCC AAGTAGCTAC  
10301 ATGGACCTGC CACCATAAGC CCTCTCCTGT CTTATGCTGT TGTAGAGGGT  
10351 CCAGGGCTCA CTCTCCAC TTGGCCCTGA GTACCTCTCC TTGAAAGGAT  
10401 GTCAGGGGCT GGGCGCAGTG GCTCACGTCT GTAACCCAG CACTTTGGGA  
10451 GGCTGAGGCG GCGGATCAC CAGGTCAGGA GATCGAGACC ATCCTGGCTA  
10501 ACATGGTGAA CCCCCGTCT CTAATAAAAA TACAAAAAAT AAAAATAGCC  
10551 ATTTGTGGTG GCAGGTGCC TTAGTCCCAG CTACTCGGGA GGCTGAGGCA  
10601 GGAAAATGGC ATGAACCCAG AAGGCAGAGC TTGCAGTGAG CCGAGATCGC  
10651 GCCACTGCAC TCCAGCCTGG GCAACAGAGC AAGACTCCGT CTCAAAAAG  
10701 CAAGCAAGAA AGAAAGGATA TCGGTTACCT GTTTCAGACA GGAATGCTGA  
10751 GACCAGGGAA AGGGGAGACT TGTGCGGTGC CTCAGGGAAC CAGTATCTGA  
10801 GCTGGGGGCT GAGAGCTCTG TGTGGGTGGA CTCTGTCTCT CCAGTCGCTG  
10851 CTGAGTCCCT CTCTTCTTTT CCCGCTGTCT GACCAACAGG GTTTTGTGTTG  
10901 GCCTGACCTC CAGTGTGAGG AACGGAACCA GGCAAGAGGC TTGTCCAGTC  
10951 AGCTCTGGCC CCAAGTTTGGC ATTCATCACT TGTTCCCTAA CCTGGAACCC  
11001 GTCCCTTCT CACTCTGGGG CTCAGCAGCT GCTCATCTAT AAAGTGGGGC  
11051 ATTTGGGGGT TGCAAAGTCA GTCATCTACA ATGCCAGGCA AGAACATGGT  
11101 TGCGTACATG TGGTGAGGTA TGAGACGAGA TCACTTTTCC AACATTCTGG  
11151 TTTTCCCTTT TTTTTTTTTT TTTGAAACAG AGTCTTGCTC TGTTACCCAG  
11201 ACTGGAGTGA ACTGGCATGA TCTCATCTTA CTGCAACCCC TGCCTCCAG  
11251 GTTCAAACAA TTCTCATGCC TCAACCTCTC AAGTAGCTGA GATTACAGAT  
11301 GTGTACCACA CCTGGATTTT TTTTTTTTTT TTTTGTGAT TTTTAGTAGA  
11351 GACAGGGTTT CACCATGTTG TCCAGGCTGG TCTTGAATC TTGGCCTCAT  
11401 GTGATGAACC CCCCTTGGC CTCCCAAAGT GCTGGGGTTA CAGGTGTGAG  
11451 CCACTGTGCC CAGACTAATT TTTTTTTTTT TATTGAAACA GAGTCTCACT  
11501 CTGTTGCCCA GGCTGGAGTG CACTGGCACA ATTATAGCTC ACTGTAACCA  
11551 CAAACTCCTG GGTTCAGCA ATCCTTCTGC TTCAGCCTCT TGAGTAGCTA  
11601 GGATTACAGG CACATGCCAC CATGTTGAGC TAATTTATTT TTTTAATTTG  
11651 TCGTAGAGAT GGGGTCTTGC TATGCTTCCC AAGCTGGTAT TAAACTCTTG  
11701 GCCCCAAGCA ATCCTCCTAC CCTGGCCTCC CAAAACGCTG GGAATACAGG  
11751 CATGGGCCAC TGTGCCAGCC TGGTTTTTTC TTCTTGTTCC CATTTTATTC  
11801 TCACATTTTC AGACCATGGG CTACTACTC CACTGAGCAC ATTTTGTGAG  
11851 AGTGCTCACA GCCCTGGGCC CGGTTGCTGT TTCCTGATCT CAGTCTTATC  
11901 AACTTGATCT TGCTTTGCTG TCATTTATAC ATTTTCTCAT TAGCTTCTC  
11951 CCCATTCTCT CTTTGTCTGC TTCCTTCTTC CTCTTTAAC TAACTCCTCA  
12001 CCTGCAACTG GGGGGAAGT GATTCTTGAC TGGGCTTGTG TGAAAAGTGA  
12051 TTGTAAAACA GATAGGTAAG TAGGGAATGA GGAGGGTGT TTACAAGAAA  
12101 AAAAAAATGA CTAAGATACA GGAACCCAAC CTAAAGAGGA AAAGACATAC  
12151 AGTTCAAAGG AGGCAGAAAG AAAAACATTA CAGATACTCA AATATATTGA  
12201 TAATCATAAC ACTTCTGGA AGATTAAAAA AATGCTGAAA CATGAATCCC  
12251 TTGCTAGAGA AATTACAAAG CCAAGAAAAT AGATAGGTCT GAGGATTAGG  
12301 GAGCTGTTCA GTTGCTAGGA GGAACACAAA AGCACAGACC CCAGACTACA  
12351 ATGGGTATGA AACCCTCTGC ACGCCTTTTG TTGTCCATCC CTGCCAAAG  
12401 CTGTTATGTA AAACCTCCG GGGGAATGAA TGAAATTATG TTTATACAGT  
12451 TCTTTCTATA TAAGTGCAGA AGAATCATGT TAAATAAATC TACAGGGCAG  
12501 GATTGTTAGT TTTTCTCCTT CTCAAGCAAA CTTCACTGCT GTCAGATAAC  
12551 TTCTCCATGT GTTTTTTTTT TCTCTTAGAA TGAGCCCTTC CTGAGAACCC

FIGURE 3, page 4 of 33

12601 CTCGGAATAG TAACTACACG TACCCCATCA AGCCAGCCAT TGAGGTAACC  
12651 CCTGACTCAC ATCTGCCTCT CTCAGACACA AACCATTTCC ACCTGCCAGG  
12701 GGCTCGGGTG TGGTACAGGT TTCAGAGTAT TCACTGAAGC AGAAATGTAC  
12751 TTCTTACATA CTGGGGATTG GAATGTACAG AAAAGGCTCC CGGACCACGA  
12801 AGCCCCAGGA TTGTCTTAAC ATGTTCTCAA GTTGCTTACC TGACGTCAGC  
12851 CCCCAGCAG AGGAAGTGTC TATGGATCGA TTTTCTTTGA CCTTGGCAAT  
12901 CCTGGGCTCA CAGACGTGGT TACTGCTTAG GCAGCTCAGC CTCTCAAGAG  
12951 GGAGAGGCAG CTGGTGTGAT GTGGCGTTGA CTTCTTGGA GGTGGAGGCT  
13001 GAGTGGGAGG GAACTACAAT TCTGGGGATG GGACCCAAAA GGAAGTGGAG  
13051 GCACGTTGTT CATGTTCCCTG TGGGCCCTTA GGCCTTGTTT GGTTCAGTC  
13101 AATCATTTCTA GTGCTGAGGA TTCAGAGCCC ATGGTTAATT CCATTGGATT  
13151 AACCATGTCT GTGAGCCTAG GACGGCCACT GCAAAGACGG CCTGGAGGAC  
13201 CCCGGAATAT ACCATGACTG GCAGTCAGGC CTGGTCCGGA TCAGGTCTGT  
13251 TGGTCACCAG GATGGGGTTT GACCCGAGT TTCAGTTTCA CACCTATATT  
13301 ATATCCAGTC TCATGTTAGG GGCTAGAAGG CATGCAGAGA AGTATCGAAC  
13351 ATGGTCCGGA CCAAGGGAAG TGAGAGCCCA GTAGAATTTT ACAATTATTG  
13401 AGCACATACT ATGTGCCAGA CACTATTTCCA GGAAGACAGA AATGTTAACC  
13451 AGACAGATGG ATCCCGGCC TCACGTAGCT TACAATCTAC TGAGAAAGGT  
13501 GTCTTATATA CATGGCTAGG CATGGTCATT TCAGATAGTG ATGACAGCTC  
13551 TGAGGAGCGT GATGGGGCTG GGGCAAGGGA GGCAAATTCA GGTGCACCAT  
13601 GCAGGCCAGG CCTTCTGAG GTGAGATTTA AACTGAGACA TGCATAATGA  
13651 GGAGACACTT GCTATACAGG GAGCCAGGAA CACAGTCCCA GGCAGAAGGA  
13701 CCATGGACCA CACAGGCTCA GAAGTGGGAC TGTGTTGGGT GTATTTGGGG  
13751 AAGAGAAAGA AGGTCAGAGT GGCTGGGGGC ATGAGAATGA GGTGGAGAGT  
13801 GGGGAAATG AGATCAGGAG TGCCAAGGAG CCAGATCACA CAAAGCCTGA  
13851 ATTACTGAGT AAAACCACTG GATTTCAAGT GGAGAAAGAT GGGAGGCAT  
13901 TGGCGGTCTC AGGAGAGAGT GACATGATCT GGTTCACGTC TTTCAAAGAT  
13951 CTCCTGACT GCTATGTGTA GAATGGGTTG GCCATCAGCA GGAGTGATTG  
14001 GGGAAAGACA TTTTATAAGC CAGCTGAAGA AACTAACCCA TATGAAATCA  
14051 TTAAGAATA TTGGATGCTA AGCTCTGGGG TGCAAGCAAT ACCAGATTGC  
14101 TGGCTGCGGG TTATGCTGTG TCCAGCCTCT CTGAATTTTC TCAGGCTCAC  
14151 GTTAGCCAG TGGAGGCTTG TCCTCATTGA ACCAGTGACC AAATTCCTTG  
14201 AGAATTGAAA CGTCAGCTGC ATCTTGTAAG TCAGGCATTT CTTCAATTAT  
14251 TCATTACCT ATTGGATGCC TATGTAGAGT GGGCACTGCA CTAAGTGCTC  
14301 GGTAGACAGT GGTGAGCCGA ATGGGTCTGG ATCTGCCCTC TTGGTTCTTC  
14351 AGTCTCATGC ATCTTTGCTT TTGCTGCTGG AAGAGCTAAA AATCCCAGAG  
14401 CTAGAAGGGC GTGTGTTTGT TTTAACAGCT TTCTACTCAA AGTAACCACA  
14451 GAAACAAAAT TCTGTCATCT GAGGTAACGT GAATGAGCCT AGAGGACATT  
14501 ACGTTAAGTG AAATAAGTCA GGCACAGAAA GACAAATACT ACATGTTCTC  
14551 ACCATATGCG GAAGCTTAAG AAGTTGACTT CACAGAAGTA GAGTATAAAT  
14601 AGTGGTTATT AGAGGCTGGG AAGGGTGGAT GGTGGTTGGG GAGTAGAGAT  
14651 AGCAGAAATT GATTAACAGA AAATTACAGC TATATAGGAA GGAGAATTTT  
14701 TAGTGTTTTA TAGCACAGTA GGGTGACTAT AGTTAACAGT TTACCATATA  
14751 TTTTCAAATA GCTAGAACAG CAGATTTTGA ATGTTCCCAA CACAAAGAAA  
14801 TGGTAAATAT TTGAAGTGAG GGATAGGCTA ATTACCCTGA TTTGATCACT  
14851 GCACATTGTA GAGATGTATC AAAATATCAC ACTATGCCTC ATAAGTATGT  
14901 ACACTTAATA TGTCAAATTA AAATAATAAA AGCAAACTA ATAAAGTGGC  
14951 CACAAAGAGG CTTTACCTGG GAGCTTTTTA GAAATGCAGA GTCCTGGGCA  
15001 CCACCCAAA CCTGCTGAAT CAGAATCTGC AGCTTAAGAT CTTCAGGGGA  
15051 TTTGGATGCA CTGATTTTGG GTGTGGTGCA TGGTTCTTCC CTTGTGACGG  
15101 ATGAGCACGT TTCAATTCCA ACCAGGATCT GTTAATCTAC ATGGAATATG  
15151 CTTATCTCTG GTTACCAAC TATCTGAGAT ATATCTCATG TGCTGATGGC  
15201 TGAATAACTT TTTACGTTGC ATTTTCTGTG AGTATTTGTC ATCTGCACAC  
15251 AAGCATGCTC TTGAGTTCAT TAAACCTTTA AACAGAAGAA ATCCATCAGA  
15301 ATGATGAATT GAGCAATCCC TTGGGAAAAA ACCAAATTCC ATAGGATTAA  
15351 GCAAATAATA TTTAAAGAA GTTCCATTTT TGCTCTCTCA TGATAGGAAT  
15401 ATTTCAACAA GTCTTATCTT CATCATCTGA CTGAACAGAT GAGATGAGTT  
15451 TTCATAGCAT CTGGCAGTCA GACTCCTGGA CAGTCAATCT GCTGGTCAAG  
15501 CCCTACTCCA TACTCAGTAT GCATATATTT GAGACTTTGG GAAGATACTC  
15551 AATTTTCCCC CAGATTCTG GTACTAATCA TTTCTATGCC CTCTGCTTCC  
15601 CATCCCACTC CTTTCCCCAG CACCTGGAAA ATATGTTCTG TATTAGAGAC  
15651 AAAGAAAATT GACTAAAAGC ATCCAGGGTT GCTTACATCA ATTTAAAAAC  
15701 ATATAAGGAA TAAGGCTGTT AAGTTAAATA TGCAAAAAGA CATACAGGTA

FIGURE 3, page 5 of 33

```

15751 TCCAGAAAAG ACAGGCAGAA ACCAGGAGCT TTACAATTTT AAAATATTTT
15801 GTGTTATTAT TCTAAAAATA TTTTAATTAT TGTCTAGGTT CTACCATTAT
15851 AATTAGTGTC AGTTAGCTTA ATTTTATAAA ACACACATAC CTGTAATCTC
15901 ATGTTAGGCA TCCAAATGCT GTGTTCTTTT GGGAGACCCA CCTGTGTAGG
15951 ACTTCATGGT TTTCTTCCCT GCTTTGGGGC AGCCACTGGC TCCATTCAA
16001 GCATAGATAT ATGGGGATAA GAAAGGTTGT GTGTGGGTGC ACATGTGGAG
16051 ACATGCACATA TGGGTTGTGC ATAGGGGTAG CTAGACACAC CCATTTCTCC
16101 CCCTTTAATT TCCCTCCTAG CCCACCTATA ACTCACAGTT CTTTCCCTCA
16151 CATGATCCTG TATGGTGACT CATTCTAGC CTCCATCAA AATCCCTTAG
16201 CTGGTTCTTC TTGGGCTGAA GCTTATCTCC CTGCACAATG AGTGTGGGC
16251 ACTGAATCTT TTCTCCTGTT GATTTAGAAC TGGGCGAGTG ACTTCCTGTG
16301 TACAGAGTGG AAGGCTTCCA ATAGTGTTC AACCTCTGGT GAGTGAAAAC
16351 ATCATCATCT CCTTCAATTA AGGGCCTTGC CGAATATCAG GTTGTGGGGA
16401 GACCCTGCAA ACATACCCTG GAGCTTAAAG CAGGACTTGC TAATCCCCT
16451 GCAGTGACAG CCTAGATCCT GCGGCCGTGCC GCCACAGCTG GGCTTCCATG
16501 TGGAGGTGCA CAGAGCTCTC CATTGGATGC TACTTCTTGT CTCCTTATAG
16551 TCCCAGTGGC AGTCCCTTAG GCCTCCCTGC CCAGTGAGGC AGGTAGAGTC
16601 AGGGATTGGG ATCTACCTGC CTGTGCTACA TGACCCTGCA GCTGGAACCT
16651 TCCTGGACCA CCCCATGTC AATCAGGCTC TTCTGAGGGT GGATGATAGC
16701 CATGAAACCC ATTCCCTGCA GTGCCTTGGT TGGTCTGAAT GAATGGGAGG
16751 GGCAAAACTG CTAAAGCCTT AAGCTGAAAA TAAGTACAAT GGGGAGCAGT
16801 GGGACAGAGT TATAGACTTC TGGTAAATG TGTACTTTAA GAGGTAGATA
16851 CCCCAGCCC CCACAACCAC CTCTCTGCTT GTCTCCCTTA GTCCACCAGC
16901 TCGGACCAGC AGACATCAAA GTGGTGGCCG CCCTGGGTGA CTCTCTGACT
16951 GTGAGTAGTG AGCCATGAAC CAGGATGGGC AGCTCAGAGT CCAGCCAGGC
17001 CTGCGCAGAG ATCTGTGCTT CCCCAGCATT GGCTCCGCTT TCAGTGCTGA
17051 GCCCGTGTTA CTGAGGGCCT ACCCATGTCA GGCAC TGAAA CACAGCCAGG
17101 AGATGTAGAA TGCCCTGTCT CGCCACCTTC CCAGTTCTGC TCAAAGCCCC
17151 CTCGTCCATG AGGCCTCCCC TCAATCCCC AGGGAGAAGC AATCCCCGCC
17201 TTCCCCACTG TTCACAGGCG TTTTGTGTGT GTGTGATGGC ACTCATCAA
17251 GTCTGCCTGC CTTTCATCAGG GGACTCATCT CCATCTACCC AGACTCAGAG
17301 TGGCAGGTCT TACACACACA CTGCCCCATG CTCCCTACTC CATTTAAGGA
17351 CATGTGCTTT GGGGCAGAGG GAGCCCGGTT CCTCACACAT AGCAGTCTCT
17401 TGCTAAGTGA ATTGTGTTTC CCAATTACTT AGCCATTGTT GTGTACACCA
17451 AACTCTATT AGCAATTCTA AGGGAAATGA GGTATGAAAC ACAGTCATAG
17501 CCCCCAGCA ACCTGTCTGG CTGGA AAAAC AAGAAACGTA CACAGAAAGA
17551 AATGCATAGT CACATAGATG ACATATAGGA CTTGGATGTT TTATTTTAT
17601 TTTTAACTT CTAAGTTTCT GGGTACATGT GCAGGTTTGT TACACAGGTA
17651 AACTTGTGGC ATGNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
18801 NNNNNNNNNN NNNNNNNNCT CCAAGTTTAA TATGGCCTGG AGAAAAGCCG
18851 GTACTATTTT TAGAAAAGGC AAATCCAGGT CCTAGCTGCT ACCCAGGGC

```

FIGURE 3, page 6 of 33

```

18901 CAGAGGAAGG CTCTTCCAGT GCCCTCAGCC TATACCCCAG CCCTGAACTT
18951 TCTTTTTTGCT TTTTACAGAC AGCAGTGGGA GCTCGACCAA ACAACTCCAG
19001 TGACCTACCC ACATCTTTGA GGGGACTCTC TTGGAGGTGA GGATGTTCTT
19051 GATGCATGCT CTATTGATGA TGCTCTCTCA GAGAGGTGTG AGTAGTGTGT
19101 TTCCTGTCAC CCCTCCAGGG ATGCAGTTGG GTCCCCAGGT CCCAGCGCTG
19151 AGACAGGAGA CTCAATGCTT GCATTACCCC TGAGGGTGAT GGGAGAGACG
19201 CCCCAGGGGC CCAGAACCCG GTTCCGGTTC TGGCTTGTGC ATATGTTGAC
19251 ACAGGGAGCA GCATGTTGGT GTGAGTTTAA CAAATATGCT TTCTCCTCCC
19301 CAGCATTGGA GGGGATGGGA ACTTGGAGAC TCACACCACA CTGCCCAGTA
19351 AGTAGCAGCC CAGAGAGGCA CCATCACTGT GGCCGTCTCT CCTGGGGCCA
19401 GGGCCTTCCT GCTGGAGGAG GGAAGAGGA GGTATCTGCT AAGAAGGGAA
19451 GTCAGCCAGC CCTGAAAAGC CCCAGACTTC CTGTGTCCCA CCCATGTCCC
19501 CACCCTGCAT GCTCATCTCA GTTACTGTGA GGGTCCTGCA GGCTCTCACC
19551 TGTGCTCTTC TCCTCCTCCT CCTCCTCTAA AGACATTCTG AAGAAGTTCA
19601 ACCCTTACCT CCTTGGCTTC TCTACCAGCA CCTGGGAGGG GACAGCAGGA
19651 CTAAATGTGG CAGCGGAAGG GGCCAGAGCT AGGTGAGTAG ATGCCGTACA
19701 GGAGGGCGAG NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
20951 NNNNNNNNNN NNNAAAATCA AAAATTAGCT GGGTGTGATG GTACACGCCT
21001 GTAATCCAG CTACTCGGGA GGCTGAGGCA GGAGAATCAC TTGAACCCAG
21051 GAAGTGGAGG CTGCAGTGAG CCAAGATCGT GCCATTGCAC TCCAGCCCGG
21101 ATAACAAGAA TGAATCTCCA TCTCCAAAAA TAATAATAAT TAAAAATAAA
21151 TAAAGATAC AAAGGAATCA AAAGATGAAC TCCCTGGCCA CGAAGAGCTT
21201 GCACTCTAGG TAAGGAGGCT AAACAAATGG GAAATAACTT TCTGAAAAAG
21251 ACAATGCTGG GTATGGCAAC AATGCAGTGC TTCGCATGGA GTACAATTAA
21301 GAGAACAGAA GAGCACACAG TATGAAGTGC ACTGTCTAAA GACAGATGCA
21351 GACCCAGAAG GGACCCTGA AATCATCCAG TCCAACCTCT TCCTTTAAAA
21401 GATGGGAAAA GTCAATCCTA GCAAGATTCA GCAACTTGTA CAAGCTCAAC
21451 AGCAAGTTGG TAGCAGAGCT GAAAGTAGAA CCACTGGTCC CTGGGGTAAA
21501 AAAGGAAATG CAAGATGTGT GGATCAGGGA GCCCAGAGAG GAGGCTCAAG
21551 GGAAAGTAGG ACTTGGTCTG GGCCTGAAGG ATGGGAAGAA GATGGCTAGG
21601 AAGAGGGGAA GAAGCGGCAT TTGTAACCTC CCTCCTACC CACGAGGGCT
21651 TATTGCCCAT GGATTCTCTT AGTCACACCT TGAACCTGTT AAAAGGTTAA
21701 AGGCACTTCT GTGGTCACCT TTGACCAGAA AAGTCTTTCT TTATAGCTTT
21751 CTGGTATACT CATCAATAGC AATAATGTAT GGGATACAAT CCTAGATCTG
21801 TAAATTCTCC TTAATGAGAA ACAAGGGTAG GGATGGTACC ATGTGTGTTT
21851 GGCCACGTAC CTAGCTTACC GTGGCCACCC AAAGATTTTC AGTGGCCAGC
21901 TCGCACTGGT TGCTGCTTTT ATGGCTTCTT CCATGGACGC TTTCAATTGGC
21951 TATATCCCTT TTGCTGACCT AACTTCTCAC AGACATTTCT TTAAACACAG
22001 CTTTATTGAG GTATAAATGA CATGCAATAA ACGTCACATG TTAAAGGAT

```

FIGURE 3, page 7 of 33



22051 ATAATTTGGC TGGGTGCGGT GGCTCACGCC TCTAATCCCA GCACTTTGGG  
22101 AGGCCAAGGT AGGCAGATTG CCCAGGAGTT GGAGACCAGT CTGAGCAACT  
22151 TGGTGAAACC CTGTCTCTAC CAAAAATACA AAAAATTAGC CGGGCATGGT  
22201 GGCATGGACC TGTAGTCCCA GCTACTCGGA AGGCTGAGAT GGGAGGATCA  
22251 CTTGAGCTCA GAGGGGTTGA GGCTGCAGTG AGCCGTGATC ACACCACTAC  
22301 ACTCCAGCCT GGGCAACAGA GCAAGACCCT GTCTCAAAAA GGATACAATT  
22351 TAACATTGTA CCTGTGAAAT CATCACCACA ATCAAGATGA AAAATGTGTT  
22401 TATCACCCAC AGGAGTTTTC TCAGGCCCTT TGGTAATCTC TCCCTCCTGC  
22451 TCCTTCCTGT CCCTACCTCA CACCCAGGC AACCATAAC CTCTTTCCA  
22501 TCACAATAGA TTAGTTTGCA TTTTAAAAA TTTTATATAA ATGGGATCAA  
22551 AGAGTATATA CTTTTATCT GACTTATTTA GCAAAATGAT TTTGCGATGC  
22601 ATCCATGTTA TTCGGTATAC CAATAGTTCG TCCCTTTTTA TGGCTGAGTG  
22651 TAGTGTTCCG TTGGCATTCA TATCGCTCAT CCAGAACACC AAATGGTATT  
22701 GTTTTATTTA TGGCAGACAT CAGGGGATGA AGGGAGAAGT AATCCTGTCC  
22751 ATCCTGGTTT ATTGGAGAGG GAGAAAAAAA AAAGTGAGGA GATGGGGAAT  
22801 GGTGCGGAAA TCTAAGTAAC CACAGAAAAG AAAACAAAAA GGATTAAAGG  
22851 AGCAGAGAGC AGGGCTTAGA AGTAAAGGTT AAAGGAGTCA TTAAGCCTGG  
22901 AAAGGAGAAA ACTGAGGGAT AATTGTGAGC TGTGACTTTT CTCAAATATA  
22951 CAAAAGGTTA TTTTAAAAAC AGGCAACTGA AGAAGAAATG AACAGGCTTG  
23001 GCTTACGAAG AAAGAGCTTG AGGAAGTATA AGGGAAAGTC CCTGAGGGGA  
23051 GGCTTGACGG GATCCCAACC CGAGTGGCCG ATGAGACTAT TGGGTGGCAG  
23101 GGGCTAGATC AATGTGGCTC CAGGGTCCAG GGCAGCCATG TGATTGTTAC  
23151 TAAGCTGAGA TTTCTTGAGA ATGGAATGAC CTTTGTACTG GTAACATCAT  
23201 TCTTCTTGAA ACCCTCTCT TCCTAGGCCA AAATCCCATG TCGTGAGTCC  
23251 TCCTCCTGA GCCGGCACTA ACGCCCTCT CTCTACCCC CACCTAGGGA  
23301 CATGCCAGCC CAGGCCTGGG ACCTGGTAGA GCGAATGAAA AACAGCCCCG  
23351 TGAGTACAGG CCCCAGGCC ACCCTGAAA GGTGCCCATC TCCTGCTGGC  
23401 TGGGGAGGGG ACAGCCCCAT AAGGGTCCCT CTCACCACAG CACTTCCTGC  
23451 TTTGGGCTAG CCAAAGATC CTCGGAGAAG CAGTCCTTAC CAAGGAGCG  
23501 CCTGCCCTGG CCACACTCCT AGACGCAGGC TGTGGCACC CTCACCCCAG  
23551 GGCCGGCTGC GGGAGGGCAA GGTGGAACAG GGAGTTGGCT GAGGTGGTGG  
23601 CCTTGGCCTC TGACAGCTTC CTGCTTTAAC CAAGAGGTGG CTTCCAGAG  
23651 CCCTATTATG TAAATGCAAG GTTCTAAAA TAGGCTTCTC ATTCCAATCC  
23701 AGTTCTGCCT CTTCCCTC ACCCTGCCCT TCTGAACTT CTCAC TAGCA  
23751 CTTTTTTTTT TAACCGTTCA GTGTTTATGC CTAGGAATTC AGCTCCCGT  
23801 GGGATTCTTA TTATGGAGGT GGCCAAGTGG AAAGCCAACT GCTTAGAGGG  
23851 CCTCCAGGCC CCAACCCGC TTCTCAGTCC ACGCTGGGCT CTTCTCCAG  
23901 TCTCCTTCCC CCGACCCTAA GAACTCATCC CAGGGGCAGC TTAGGCCTT  
23951 TGCTTCTAGC TGCACTCTT GCCTACAGCT CCCTGGAAGG CTTCTATTG  
24001 GGGGGACGTG GTAATCCCT CGGCATTTAA TGGGCCAAGG ATATGTGGGA  
24051 CACATCCACA TTCTACTTCT CCAGGGACAC AACTTTCTTA AGATTCAAG  
24101 GGGAAAATAG CCCTCCCTTG TGAAGCAGA ACCCCGTCCC CCGCCAGCGC  
24151 CCACCGCCAA AAAAAAATA CATCCTCTCT GTGGAGCACC TTATCCTAGC  
24201 ACCAATTGAG GGCTGGGAAG CCCCACCTTG TTGCTTTTCT TTTTTTTTTT  
24251 TTTTATAGAC GAGTCTCGCT CTGTCACCCA GGCTGGAGTG CAGTGGTGCG  
24301 ATCTCAGCTC ACTGCAGCAT CTGCCTCCTG GGTTCAGTG ATTCTCCTGC  
24351 CTCAGCTCC CAAGTAGCTG GGATTACAGG CACCTGCCAC CAGGCCAGGC  
24401 TAATTTTGT ATTTTATAGT GAGACGGGGT TTCATCATGT TGGCCAGGCT  
24451 GGTCTCGAAC TCCTGACCTC AGGTGATCCA CTTGCCTCAG CCTCCAAAG  
24501 TGCTGGGATT ACAGGTGTGA GCCACCGTGC CTGGCCCAAT TTGCTGCTTT  
24551 TCTCTGTTAC AGTATAAATA AGACAAAAG CACTTGAGAG GCGGGCGGGC  
24601 TGGCCAGTA GCAACATTTG TATGTGCCTC CCACCAAGGC CTAACTCAG  
24651 GATCTTCTGT CCCCTCAGGA CATCAACCTG GAGAAAGACT GGAAGCTGGT  
24701 CACACTCTTC ATTGGGGTCA ACGACTTGTG TCATTACTGT GAGAATCCGG  
24751 TAGGCCCCCG ACCAACCCCA TGGGGACCTG AGAAGGAAGG TGCTGACCTC  
24801 TGGCAACACC CTTGCCCATC CATCCCTGGC CTTGCCCGA GCTCCTCGCT  
24851 CATGGGAACC ACATTTGCCT GCTGCCCCAG GCCCTCCCTG GTTTACACAT  
24901 GCCAGGCAAG GCCCAGCCTT TTCTACTGCC TGAGCGACCC CTGGAAGAGC  
24951 AGGTGCATTG GTTCCCAAT TCCAGAAGTA AGGCCAAGGT GGACCCACTG  
25001 TAGGCACTGC TGAGGTGAGG CCTCTCTTAT CCACACAAAT ATGACCTCTG  
25051 GTACCAGATA GGGGACTAGC CATCCTCACC CCATCCCTGC CCTGTTTCAT  
25101 TTTGAGGAAG GGCAAAACAA TGTCTAAAT GGGGTGGAT GGGTCATCAC  
25151 GAGTTAACCA AACCTCAGTG GTGGCCCTGG GAGCCCAAAC CTGTTCTCTGA

FIGURE 3, page 8 of 33

```

25201 TGTTCCTATG GGGAAATATG TCCAGACTCA CAAACTTCTG GAAGTTGTAA
25251 CTTCCAAAAG TTTTATTTT GGAATACATC TTGTTCTAAG TTGGGGACAT
25301 TCTGTATTTT ATTTGGCCGG GAAAAGGCAG CTGCCCAGCC TCAGAGATTG
25351 TGGTCGAATG TTTGACAGTC ACAGATGGTG GTGTGGGGGA AAAGGCATTT
25401 ATCACTCCTT GTGGCTCACA GAGGCCAACT ACCAGAGGCT TGATGAGAAG
25451 TATCCTCCCA CACACAGGAG TTGGTTTTC AAACCTCTTC CCCTCAGTTC
25501 TCTCCCTGAC CACCCACACT CTAGAGGCAG AAGTGACCCT AACTCATGGG
25551 GATTTAAGTA TTGCTGTGTT CTGGACTCCG GGGATACCTG GACCCAGAG
25601 CTGCGTGGAC ACTGTGGACG CTGGCGTAGG GAAATGCCCT CTACTAGTCC
25651 CAGGATGTGT GTTCTTCATG GGAACAGTCA GGTTTATACT TCCGAGAGCG
25701 TAGTTTAGTT GAAAGGGCTG GGCTGCCCCG ACTAGGATTA ACTCAGACTG
25751 TTTTAAAAAG AAGGAAAGGG GGAAGGCAG AAACCTCTGGG AGACACGAGG
25801 TCCCCTCTCC TCTATTTTAA CTCTTCGGCA TGGATTGTCT ATCTTGTTCC
25851 TTTCCCCTTC TTCCCAACTC CCAGTACCCC TTCTGGTGGC TGTGGCCAGA
25901 AACTCAGCGA ACAGCACTTG TCCTATGCCC ATCAGTGTGC TGGAATTGAG
25951 TGCACGGTAT CTCACCTGGC CTCAGCTCTT CGTCTCCAGA AAAAATAATG
26001 GGCTGCCTGA GCTCTCCCCT CTGCTGAG TGGTGCTGCT TTGTGGGTGC
26051 CTCAACTCCC ACTTCTCTGT GACATGCTTT CTTCCATGAG TATAAGAGCT
26101 TCAGGTTACC ACCCGCACCC CCACTGGTAT CAGCCTGTGA CACCTTCTGG
26151 GCCTGTAGCC CAGAGCCACA TCTAAAATA GAGGCCATC TCCCTCTGCT
26201 ATAAAGCAAA GCCCTGAGAT TCAGCCTGCA AGGACTTACT GAGCACCTAC
26251 TATGTACCTT GTTTGCATCA CCCAGGATGC TGTGGACACA CCTCTAAATC
26301 AGCCTCCTAC TGGGGAGATG GTTCAGAGGA AGAGAACCTT ACACTGAGTC
26351 ACAGGGGATA GAAGTTAGGG GAACACAGGA GAGCAAAACA TTTCAGGCAG
26401 TGGGACCAGC ATGGACCAA GCCCAAAGGA AAAAGGAAGT GTGGCCACCC
26451 AGGGCATGGC AAGGGGCTGG AGAAGGCTGA GGTCAAGTGA CGGATGGGAC
26501 TGCCAAGAGC CAAGGCCAAA AAGTGGCAGG ACCCAGCACT GGCAGAGTCC
26551 ACTGTTGGGT CTGAGATTAT GTAGAGCAGG GTGGGGGTG GGATTGTTCA
26601 TGGTGTCTAG TAGGGGACAA GGGATGATTC CTTACAGAGA CTCAGCAGCA
26651 ACAAGAACTG GGCTTCTCAG TTTGACCAGG ACCACCGAAG CCCCTCTGTA
26701 CCCACTCAGT CATTTAGCCC AGGCCCCAGA GCCCTCCTAT GCTCTTGCCA
26751 TTCTCTCAGA GCGGGCACCA GGGGCTAAAG AGAGTACCCT TTTTCTCTTA
26801 CAGGAGGCCC ACTTGGCCAC GGAATATGTT CAGCACATCC AACAGGCCCT
26851 GGACATCCTC TCTGAGGAGG TAGGAGAGGG GTTACGTGTT CCTGGGTCCC
26901 GCCAGCCACC TCCCTGGGAT GCATGTAGGC AGGCTGTGTT CAGTGAGATG
26951 CTCACGGAGC AGAGACCCGC CATGAGTGAG CACCTGGATG GCAGGGAGGG
27001 AGGTGGCTGT CAAGCTCCTC TGCAGGGAAG ATTCTCACTT GGCCAGAGAC
27051 AGGGTTGTGT GGTAGCAATG AGCTTGCTC TGAACCAATT GGCCAGGTT
27101 TGCGCCAGC ACTGTGGCTT CGGGCAAGTG ACTTCCCTGT GTCTCAGTTT
27151 CTCAACCTAT AAAGCGGGGC CACTCAAGAA GATTCAAGTGA GATACTACAA
27201 GTTGATCCCC CTCTCTGGGC CTCAGTTTCT TCATTGGTAA AATTGAGGGG
27251 AGTGGGAATT GGATTGTAGA TGACCCCAAG GTTCTCTCCA GCAGTAGCCA
27301 GTGCCCTAAC GAAACCAACC TCCACTCCCT GCAGCTCCCA AGGGCTTTCC
27351 TCAACGTGGT GGAGGTCATG GAGCTGGCTA GCCTGTACCA GGGCCAAGGC
27401 GGGAATGTG CCATGCTGGC AGCTCAGTAA GTGGACAGGT CACCGTCCCA
27451 AGGCAAGGGC ACCTGGGGTG AGGAGGGCTT GCAGGTGCCA AAGGAGGAGA
27501 CCAGTTGAGG CAGAGCCAGG CAGGCCTGCC AGAGGGTAGA CATGGCTCAG
27551 GGGCTTGGAC AACATCAGGA AGTACCTCTA CATTTGCAA TGCCTACTGT
27601 ATGCAAGGTG CCTCATTTCT CTGGACCCCT TTTTGCTTTT CTGTGAAAGG
27651 AGACAGACCA AATGATCCTT AAGGCTCCCT GACATTGTCA GTGATTGCCA
27701 AGGCAACCT TGGCACGCTG CTCCCTGTTG AGAAGCAGCA TGGGGCCATG
27751 AGCTTTCAAG GCTGTACAT CCAGCCTTGA CTGTTCTGCC ATTTAGGAGC
27801 TATGGGACCT TGAACAAACC ACATAACCTC TATGAGCCTC TGTGTTCTCT
27851 ATCTGTAAAG TGGGGGTGAT GACACCTTCC CTGCAAGGTA GATGTGAGGT
27901 CAAGAGGAAA CAAGGTACCT GGCCTACCAA GACTACCAAG AGCAGGCTCT
27951 TAGGAAATAG CTCTTATTCC ATCCTTGATG GGCCTGTCC TTGATAGCTG
28001 GGCTTGGAGG CAAGTGCTA ATGGGCAAGA CAAGAACTCC TATCGGGGGC
28051 TGGAAGTCAT TAAAGCTCTT GAACCTGGT AGGAAGTTGC CATGTTCTGA
28101 GGGCACAGGG CCTCCACAG TTTGAGTGAT TATTGCTATG AGAGAGGAGG
28151 TTCTCCAGGG AGCTGAGGAG TCCTACACCT GGGCTCAAAT GGATTGCTG
28201 CAAAGGTGAC CAGCTGGTTC CCATTCTTGC AGGGAACCTT GCTCTCTCGC
28251 AGGGAACGGC TCCTCCAGAG TCTGTCTGTA TCGTGTTCCA TGTGTCTCAGG
28301 GTTGCTTCCA GCCGGTTGGC CCTCCCCAG TTTCCACAG ACTTCCACA

```

FIGURE 3, page 9 of 33

28351 CTGGAGCCCT GAGGGAGGGT CCTAAGCAGT TGCAGGAAGA GCTGAGAGGC  
28401 CCCCAGAACT TGAGGAGCGA TTCCAAACCC AGGGACAGAG CCATCGTGGC  
28451 TGGTTTCCTA AACTCCAGTC TCCTGTCTAC CCAGTCCTGC TCTGGAGAAA  
28501 TCCCAGGGAC CACAGGCTTG GGAAGGAGGA AGGGGAATAG GCGTTCTGTC  
28551 CACAGGGAGG TCCAGGCAAC AGCTTTCCCT CTTTCTCTAT GAACAATCAT  
28601 CCTCTGGACC TCAGGGCTCC TGAGTTAGCA TTCTGTAACC TGGGTCCAAG  
28651 AATCAGCCAA AGGTGTATTG TGGGATACT TGTGTGTAC CCCCCGCCCT  
28701 AGGTAAGGCA GCACAGGCTG CAGGCCCCCTG GGGTAGTGGC CTGCTCTGTG  
28751 TGTCAGAGCC AGCCTCCCAG GAGGACAGAG CCACAGTGCC CCAGGCAGCC  
28801 TCAATACAAC ACTCCCTGTC TCACAGGAAC AACTGCACTT GCCTCAGACA  
28851 CTCGCAAAGC TCCCTGGAGA AGCAAGAACT GAAGAAAGTG AACTGGAACC  
28901 TCCAGGTAAG CCCTGCAGCC CTTCTCTTAC TGACCCAGCT GGGGGGCCCC  
28951 CTGTACTCCA AGGACTGGGA AATCGAATGC CCAGCAGGAT GTGGCCAAGA  
29001 GCAAGCCACT CCCTAAAAGC AGATTGCAGC CCCTGAAATA CTTACCCCTG  
29051 CAAATTGAAC ACCAAGGCCA GGAAGGGAG TGAGAGACCC CAAAGTGGAA  
29101 GCTGAGAAAA TCCCCTTCTC CCAGCGGGTA GGCAGCAAGA GATTCCCAGA  
29151 GTAGACTCCT TGTGGTAGGG CCCATTCCCC ACCCAGAGCC ATGTGTAATA  
29201 ATTACTACT ACTTCCTCCC CTCCCTTCAT TAAAAACAAA AGGCTTAGGC  
29251 CCGACACAAT GGCTCACGTC TGGTGTCCCA GCTACTCAGG AGGCTGAGAT  
29301 GGGAGGACAG CTTGAGCCCA GGAGTTGGAG GCTGCTGTTA GCTATGATGA  
29351 TGCCATTGTA CTCTGCCTAG ACAACAGCGT GAGACCCAT CTCAAAAAA  
29401 AAAAAAAGAA AAAAAAGAAA AGGCTTAGCC CTGCCCTACT TAACTCTACC  
29451 TCAAATTCTC CTTGCCCTCT CTCTGCCCCC TTCCATCTCC CCACCTCCAC  
29501 TCCTGCTTAT GTCTCTGCC TCTCTGCTCC CTATTGTTCC CTCTCAGGCT CAGGTAGCAT  
29551 TTCCATTCTG CAAACTGACC CTCCTTCATT CACAAGGCAA GTCTGCTTCC  
29601 CTCCTCTAAG GAGCTTCCCC TGCTGAACT TCACCCGCGG ACATCTCCCC  
29651 ATATCACATT CAGTCTGTAC TTGATGGGCC CTAAAAGCCC CAAAGGGTTC  
29701 TCATGTTTTT ACATCTTGGC TCATTTTTCC AGATGGATGA TAACTCCTT  
29751 GAAGATAAGT ACATCTAGTC TGTTCCCTTT ACATTCCATG CTTGGGTACT  
29801 TAAATCCAGC CACCGTGGAC TCTCCTCCCG CAAAGTTCAT GGGCATTTTG  
29851 GGAGCTGGTG TTGAGATGCT CCCCATCTGA CCTGCAGCCC CATGTTCTAA  
29901 TTGACCTCTT CGTGCAGTGA GAGGAGGGGA GGACTTTGGC CTATGCAATC  
29951 TGGTCAGTGG CTCAGACCCA GCCTTTTCAGG CAGAGGCTTT GGAATGGGAC  
30001 TGGGTGGAGC TGTGTAGCTA GGGAGCTTCT CCCACCAGGA GCCGCTGGGT  
30051 TCAACTCATC TCTGATCCTG AGAACCAGCA TAGGGCTTTG AAATGTCCGT  
30101 GCCCATGAAT GGGTGGAGAA TAAAAGTATG TTTGCATCCC ACTAGAGTAG  
30151 CCCCTTAAAG TCACTGTCTT TTAGGGTGAG TTGACTCCCG TCAACAACCA  
30201 ATCCAAGGCA GCAGGACTGG ACCCTGTCTG TGCAGCCTTG CCAGGAGGGT  
30251 TGAGCAGCTT CTCTCTCTGT CCCCAGCATG GCATCTCCAG TTTCTCTAC  
30301 TGGCACCAAT ACACACAGCG TGAGGACTTT GCGGTTGTGG TGCAGCCTTT  
30351 CTTCCAAAAC AACTCACCCT CACTGAACGA GGTGAGCTGC AGGTATTTTA  
30401 GGGAGGCTCA CGTATGGGGG CCTTATCACA GACGATGGAT GTATTTCCCT  
30451 CTCTAAGTGG GCTTTTTTTT TTTTAAACC ATCTCTCTCC AAGAGGATTC  
30501 CTGAGGTGG CTTTTTCCAC ATTACCTCCT TTTTGTGGGG GCTGGGCTGT  
30551 GATTGGAAC CAGATGTACT TTGAAAGGAA ATCAATAGTG ACTAAGCTCC  
30601 CAGGCCCTGG CCTGATGTTT TCTGGATTGG GATAGAATGG AAAGCTTCCT  
30651 AAAAATGTTA CTCTTTTCAA CTCTTAGGAT AGGGGTGCTG AAAGAAAAGG  
30701 GAGAGACTAT GGGTGGGTCC AATTCTTGTC TGTTTAAAAA GAAAATTCCG  
30751 GCCGGGTGCA GTGGCTCATG CCTGTAATCT CAGCCTTTGG GAAGCCAAGG  
30801 CGGTTGAATC ACGAGTTAG GAGTTTGAGA CCAGCCTGGC CAACATGGTG  
30851 AAACCCCGTT TCTACTAAAA ATACAAAAG TTAGCTGGGC GTGGTGGCAG  
30901 GCACCTGTAA TCCCAGGTAC TCGGGAGGCA GAAGTTGCAG TGAGCTGAGA  
30951 TTATGCCACT GCACTCCAGC CTGGCTGACA GTGCGAAACT CCGTCTCAA  
31001 AAAAAAGAAA AAAGAAAAAA AGAAATTCTA AATTCTGGGA GTTTTTCCAT  
31051 CAGTATCTGA GCAAGTTGGC AGGAAAGTTG AAAGAATGAA AGGAGACATG  
31101 CCCAGGGCAC CTGCTGGGAG AGTGAGTGGG GCTCAGGTAG CAGAGCCCTT  
31151 TCCCAGGATG ATAACCTCCT TGCCGTTGGT TGCAGAGAGG GGACACTGAC  
31201 CTCACCTTCT TCTCCGAGGA CTGTTTTTAC TTCTCAGACC GCGGGCATGC  
31251 CGAGATGGCC ATCGCACTCT GGAACAACAT GGTGAGCAGC CAAGGGCCTG  
31301 GTGGGCCCTT TCAAGGGGGG ATCTAAGGAT ATTGACACTC TGTCTCACAA  
31351 TGGCAAAACT ACTGGAGACA TGGCTCCTTT CTCCCCAAAG CCCAAAGTGG  
31401 CAGCACACCT TATTGGTCCT GATAGATTAA TTCCAAAGGG AAAATACCTT  
31451 ATATTTATCC AACACCCTTT GAAAGTTATA CAAACACACA CTCACACAAC

FIGURE 3, page 10 of 33

31501 TTTATTCTTT GTTCCTTCAG CAATGCCAG GTACTGCGAG GGGATCCCTT  
31551 TGTAATCAGA TAGGTTGGCT AGATGAAAAT ACCAACTTCT ACCTCGTACT  
31601 GTGTGACCTT GGGCAAACGA TCTCTCTGGC CACCTGTATC AACATCTATA  
31651 AAACAGTGAA AACAAAGACAG GTCTCAGACA ACGCATTGAG ATCATGTGTA  
31701 CATGGCACCT AGCACAATAG TTAGCACTCA GCAAAATGTCA CCACCATCAG  
31751 CCTTCCAAGC ACTCCGGGCT CAACTCATAC CCAACTCAT TCTCTAAACA  
31801 TCGAAAAGTG GAGATCCACA CAGCCTGTTT TCCGAGGCTG ATACCTATTC  
31851 CAGTCCTTTC TGATGGGAAG AAGGGACCTT ATGAAATGAA CATACAGTCT  
31901 GGGGGTCTTT CAGGGACACC TGCCTGGTGC TTCCACTCTG CCTTCTGTGG  
31951 CTGGCCACCA GCAACTGAAC GGTTCCTGCA CAGCACTTGA CCTGTCACCC  
32001 CCAACAACCTG GATCCTCTTG CACGGAGCAA ATGAAATGCC TTCCCAACCC  
32051 AATGGTTTCT TTTAATCCAG GCTCAGTGGG TAACACAATC CCCACCCCAA  
32101 CCTGTATGTT CCCTCCTTTG TCCTATGACA ACTAAACAAG CTACATTCCA  
32151 GCTCCTTTTA TCACAGTTTC AGGCCCGTAG TGTCTCTGCC AACCACCGCT  
32201 GTGCAACCGT TCCCACCCCT GTCAGCTCAT CCAGTATGTC CAGCATCCCA  
32251 CTCGGCTGAC TCACAATATT GACTTTCTCC TTAGCTATAC CATCTCCTCC  
32301 TCTCTAGCAA CCTCTTCTTT TAAGAACAGC ATGTAAACTG GCTTTATCCT  
32351 TGGCCTAGTT AATGGCAGAC TCAGCTTATG TCGACTTCCA TTGTCTAGGG  
32401 GTTTTCCTCC TGTGGACATC ACGTACCTGC CCACTCCAAG AACTTCTATT  
32451 GTACTCTTTC AGCCCAAGAC TCCGGATTGT AACCAAAATG TCTTTGAGTT  
32501 TGCCCCACGA TTTTAAATC AGTTTATATG GTATAATTCC TGTTTCTTTT  
32551 GTGGATTTTG TTTTGAAGGC GGTGTCTCTT CACTGGCTGA ATCATGTGAC  
32601 TTTATTCTTT TGTAATAATC TTCCCAAAGA AAGGGTACCT ATTCCTGTGT  
32651 CCTTTTCCCC TGAGACCTCA GGGGATTCCA CAGATGCCCT TGGCCCTTCC  
32701 TTCCAGTTT TTTTCATCAAG GTATGGCCTT CCTACCAGGT GGCCTCCAA  
32751 GTCTGCTTAA ATCTGGGACC CTCCAGGAAT CTCCTGGGGC TGGATAGCCA  
32801 TAGTGACGGC TGAACATGA AAAAGAGTCC ATTGGTTTCT TTTCTGTGTA  
32851 ATTAACAATG TAGCTCTGGC CAGGCACGGT GGCTCATGCC TGTAAATCCCA  
32901 GCACTTTGGG AGGCCGAGGC AGGTGGATCG CTTGAGCCCA GGAATTAGAC  
32951 ACCAACCTGG GCAACACAGG GGAGATTCTG TCTCTACAAA AATAATCAAA  
33001 ATATTAGCCA GGTGTGGTGG TGCATGCCTG TAGTCCCAGC TGCTCAGAAG  
33051 GCTGACGTGA GAAGATCACT TGAGCATGGG AGGTCAAGGC TGCAATGAGC  
33101 CGAGATGGCA CCACCGCACT CCAGCCTGGG CAATAGAGTG AGACCTATA  
33151 TCTCAAAAAA CAAATAGAAA AAAAAAATAT ATGTAGCTCT GGCCTTCTCT  
33201 TCTAAAGCAG TTCAGTAGCT CTTCCCATT CACCCAGTAA GAGGCCTTTA  
33251 TTTCATAAAG ATAAGTGGGA GGAGTTTAGA TATGAAAACA AAACGTAAAC  
33301 ACCGCACTGG AGCTATTGTG GAAACAAAAC AAGACTGTCC ATGGTTCCCC  
33351 AGCCATTATT ATCTCAGCCA TACCCCGAAT TTCAAAATAA CAAAAACAAA  
33401 ACTAAAGCCA TCCAGGGGTT TCTTATCCTA GGCTCTATAA TTTGGGTAAA  
33451 TAATTATACA GTCTAATGTT TTCATCCAAA GCCAATCTTA GACATAAAGC  
33501 TGTAGCATGA TGCCAACTTT TCAGATCGGC TTCTGGCTGG AATTTCAACC  
33551 CTAGAGTAAC AAAAAATAAA TAATAGACCA TTAGAGCTGG AACAGACTGA  
33601 GAGGTCATCT AGCCAGAACA TTCTGTAAC TAAAGCATAGA AACATGAAGC  
33651 AGTTTGCCCA ACATAACACA GACTGTTTCT GGCACAAGGG GGATTACAGA  
33701 CCAGGTTTTT CTAGTCCTTT CCTGGTGACC TGGGCATGCC ACCACCCTCC  
33751 CCACTGCTCC CAACCTGATA AGCACATATA TACCCGGTGA ATTCATGTCT  
33801 CACAATTAGA GTCCTATGAC ATAGTGTCTG CAGGCTTTGG CTGATGTTCC  
33851 CATAGTGTCT GCAGGCTTTG GCTGATGTTT CCAGGGTTCC CTAAGTAGAA  
33901 GCAAAAAGCA CCTTAAACTA TTTCACTTTA TTTCACTCTC TGCCCCCTCT  
33951 CTCACGTCCT TCTCGAGACT TTGCAAAGG CAAAGCCAGA AGCTCCAGCA  
34001 GCACCAGGGG ATATTTTCTT CTTCTCTGCT CTTCTTCTGT CTTCTTATCT  
34051 GAAGAAGTTT CTCTTTCCCG AGGCCTAGTC CTCTACTGCT GCCTCTACTC  
34101 CCTCTTCTGC AGAAATCCTG CTCTCAGCCA GTGTTTGTAT CTCCCCAGGT  
34151 GCTGGGTGAC AGCTCCAGCC TCCTAACTGA CATCCCTGTC TTCAGACTTA  
34201 GAGCTCTTAG AATCGTGACT CTCAGCTCTG GCTGCATATT AGAATCATTC  
34251 AGGGACATTG TGTATGTGTG TATGTATGTG TATATATGTA TGAATGTGTG  
34301 TGTATGTGTG TGTGTGTATG TATGTATGTG TATGTGTGTA TGTATGTATG  
34351 TATGTATGAC AGAGTCTCAC TCTGTTGCCC AGGTTGGAGA GCAATGGCAC  
34401 CATCTCAGTT CACTGCAACC TCCGTCTCCT GGATTCAAGC GATTCTCCTG  
34451 CCTCAGTCTC CCAAGTAGCT GGGGTATAG GTGCATGCCA CCATGACCAG  
34501 CTAATTTTTG TATTTTATAGT AGAGACAAGG TTTGCGCATG TTGGCCAGGC  
34551 CAGGCTGGTC TTTAACTCCT GACCTCAGGA GATCCACCCA CCTCGGCCTC  
34601 CCAAGTGCT GGGATTATAG GTGTGAGCCA CAATGCTTGG CCATCCAGGG

FIGURE 3, page 11 of 33

34651 ACTTTTAAAA CAATTAGTGC CTACAGCCAC TTTGGAAAAT TCTTTGGTAT  
34701 ATTTAATAAT GCTGAACACA TGTATTTCCT GTGATCCAAG GATTTCACCTC  
34751 CTAGGTATAT CCCCCAAAAG AAGGTATATA TGTGTCTACC AAAAGATACA  
34801 CACAAAAATA TTCACAGCAG CACTATTTAT AAATAGCCCC AAAGTAGAAA  
34851 CTCCCCAAAT GTCCACTGAC AGTAGGATGG GTGAGTAAAC TGTGGCACAT  
34901 TCATACTAAG GAATACCATA CAGGTCTGAC CGCATCTGTG GCTTTAAACA  
34951 AAAATCAAGC AGGGTGATGT GACACAGAGT AATGGCTGGG AAGAGGGAGG  
35001 CCTCACTGAA GAAGTGACAG CTGAACAAAC TTCAACAACA TACAATAATA  
35051 TCTATAAAGT TCAAAAGCAA GCAGCTTGGC ATATGGGGTT AGACGTCAGC  
35101 ATGATGGTGT AGAGACTCAC TGGGGGATGA ATAGTCCTGG AAGAAGGTGG  
35151 AAAGGGGCTT TCGAGGACTA TAATAGTCTG TTGCCTGACT GGATGCTGGT  
35201 ATGTTTCATTT TATCGAAACT TATCTGTTGC TCACCTATGA TTTGTAATCG  
35251 TTTCTATGTG TATGTTAGCT TCAATTAAAA GTTTACTTGA GGCCGGGTAC  
35301 AGTGGCTCAC ACCTGTAATC CCAGCAGCTT GGGAGGCCGA GGCAGGCAGA  
35351 TCCCCTGAGG TCAGGAGTTC AATACCAGCC TAGCCAACAT GATGAAACCC  
35401 CATCTCTACT AAAAATACAA AATTAGCCAA GCGTGGTGGC ACGTGCCTAT  
35451 AATTCCAGCT ACTTGGGAGG CTGAGACAGG AAAATCGCTT GAAACCAGGA  
35501 GGCAGGGGTT GCAGTGAGCC AAGATTGCAT CATTGCACTC CAGCCTGGGT  
35551 GACAAGAGTA AAACCTCTGC TCAAATTTAA AAAAAAAAAA AAAAAAAGT  
35601 TTAAGTGAAG AACAAATATCA GTGCCTGACC GGGCTTATCC CCAGAGAGTC  
35651 TGACTTAATT GGTCTGGAGT GCGAGCTGGA TTCGGTACTT TGTGAAAGCT  
35701 CCGAGATTA TTTAATGTG CAGGGTTTAT GAACCGCTGC CTTAGATCTG  
35751 GTCCCCACAG AGAAATCAAG TAATCTGTAT AAAAGAAAAC CTGACCCAGT  
35801 CACTCCCCTG CTTTCAAAC TCCAAAGCCT CCCACCTCTG AAGGAGGCAG  
35851 GCCAGGCCCC ATAGCACAGC ACACTAGGCC TCTGGGACTT GGCCTGGTTC  
35901 ACCTGATTAA CCTCTCTGGC TACCATTTC ACCAGCGTCT GCCTCGCATG  
35951 TTACAGTCTA GTGACTCCAG CAGCGTCCTG CACCACCTGT GGTGTCCAC  
36001 ACCTCTGCTA ACTCTTGCTC TCCTCCTTCT CCTGGATTGC CCTTCTCACC  
36051 TCCTTGCCCA CTCCACCCT CAACTCAGGT GCCACCTCCT GCAGGAAGCT  
36101 ACCTCTGAAT CTCCAGGACA GGCCAGTGGC CCACCCAGGT CCATTACACC  
36151 CTGCCCAGTC CTGTCAATTTG CTACGTGGTT GGTAGCCACA GTGCCTGGCT  
36201 TAGGAAAGAC TGGTCTAGG AAAAACAATT TCATTCCCTG TGGCCAGCTC  
36251 CAAGCCTTCC CCCGCCAAGC TTCTCCATTC AGGTCTCTGT GAATTTAATT  
36301 AATTCAATCA TCCATCAAAC AAGTATTTAC TGAGCACTAA TATGTGCTAG  
36351 GTACTGCTCC AGGTGCTGAG GACTCAGCAG TGAAAAGATG ACTGCTACTC  
36401 TCATGGGACA TACAGGATAG TAGGGAAAAG ACAGATAATC AACAAGGTCA  
36451 TTTCTGACCA CATCTGTGGT TTAAGAAAAA GTCAAGCAGA GTGATGTGAT  
36501 ACAGAGTAAT GGTGGGGGAG AGGGAGGCCT CCCTGAAGAA GTGACAGTGA  
36551 ATTGAGAAGC GCATGTCAAG GGGTTGCCAG GCAGAGGAAA TAGGACCCAC  
36601 ATGGGCCTAG AGTCAGGAGT GAGCTTGAAG TGTCTGAGGA ACTTAAAGGC  
36651 CAATGTGACC AGAGGGAAGT GAACAAGGTG AAAAAAGTTGG GCAGGGGCCA  
36701 GGTCCCCTAG TGCTTCTAAG CAGTAGAGTG ATATGCTCTG GCTTACCCCT  
36751 GGGTCCGTGT ACCCTGGACT GGAAGAAAGC AAGGGTGGAC CTGGAAGAGC  
36801 CACTAGGAGG CTGCTGTGTA TGGGTGAGAG AGGAAGGGGG CTGAGAGTAG  
36851 GGTCAGGGCA GAGGAGGAGA GACGCTGTCTG TGGGCTGGCG GATGGATGAT  
36901 GGGGAAGAGG AACAAAGGAT GACTTTTTGG TTTGGGGTCT AAGAAACTGG  
36951 GTGGATGATT GAGCAGGTAG AGAAAAAATC AGCGTGGGAG GAAAAAAAT  
37001 CAAGACTTCT GTTTTGGACA TGGTGCAAAC TGCCTTCCAG ACATCCACAT  
37051 AGAGGTATCA GGATACAGAA GTTTGGAAC CACAGAGGAA GTCAAGGCTG  
37101 GAGATTGAAA AAAAAAAAAA AAAAAAAAAA AGTGGGGTTA TTAGCATAGA  
37151 GGGCCAATAT GGTGAAACCC TGTCTCTACT GAAAATACAA AAATTATCCA  
37201 GGCATGGTGG CATGCACCTG TAATCCCAGC TACTCAGGGA GGCTGAGGCA  
37251 GGAGAATTGC TTGAACCCAG AGATGGGGTG GAGGTTCAG TAAGCTGAGA  
37301 TCGTGCCATT GCACTCCAGC CTGGGTGACA GGGCAAGATT CCATCTAAAA  
37351 AAAAAAAAAG CCACTACAGG ATCAACTAAG AGCTCCTAGA GAAAGAATAG  
37401 GTAGGTAGAA AAGAGTGTAA GGCCAACTAC CTAGCCCTGG GCATTCAATC  
37451 CAGCTTTCAA CTCCAGTGAG AGATGAGAAG GAGAGTGTGG AGGTAGATGG  
37501 GAAATGAGAA ACAATGCTGT GTCCAGAGAG CTAAGAGAAG TCAGTGTTC  
37551 AAGAGAGACA GAGCTGTCAA CTTTGATGGA TGCTTCTGAG AAGCCAAGCA  
37601 AGTTGAAGAC AAAAAAAAAA AAAATGATCT TTGGCTCTGC CCATATGGCG  
37651 ATCGTTGGTG GCCAGGGCCA GAGCTTCCAT CCAGCGATGG AACTGCAGA  
37701 CTGGCTGGAG CGAGCAGCAG AGAGAAGGAG AGATTAGGAA GTGCTGCCAG  
37751 CACCTATAGA CAGCTCTTCC CAGAAGTTAT GAGAAGTAAC AGCCACGGTC

FIGURE 3, page 12 of 33

37801 ACTGGAGGGG ACATGGATCA AAGAAAGGGC AGGTGAAGGA GGGGAGATGT  
 37851 CGGAGCAGGT TGTGTACTGA CGAGAAGGAA CCAGTAGAAA GGGAGAACT  
 37901 GATGCACTCA TCAAACCTT GTAATCACGA TCATCTTCTG TGTGAATTAG  
 37951 TTCTGGGTTT CTGGAATAGC ATCGGGAATC AGCCGCGCTG ACCTTTAGCA  
 38001 TTTATTCTGT CACTGTTACG ATAGACTTGA GTTTCCTCAG TTCTTAAGAA  
 38051 AGTGGAAATA ATACTACCTT ATGTATGTAA GCCACCCAA TCACACGTGC  
 38101 TTTCATGCCA TCTTCTCAT TGTATGCTCAC AACAAACCA GGTGTGAGGA  
 38151 CAGGATGTCA TGCCTGTCCC TAACCTCAGA GAAGTAGCCA GCCCAGCACT  
 38201 GCACAGCTCG TTACCAGCAG AGCCTGGATC CCAGCCATCT GCCCATCGTG  
 38251 CTCAGTCTAG TCACCCTAGC ATCTCTCCCA GGAACAGAAC TGCCCTCCCT  
 38301 CCTCCAATTG TGTACTAAG GAACGGGTGT AAAAGGCCTT CAGACAACAG  
 38351 AAAGTGAGAT CATGGGCCAG GTGTGGTCAT GAGCTCAGAT AGTGAACCTT  
 38401 CACCTCCCTC CCTGGCAATA CCCTGTGGTC AGGAGCAGGC AGATTACACA  
 38451 AAGAGTGGAG GCTAGACGTT CCAAACAGAC TCTGAATAGG TGACAGTGCC  
 38501 AGGGGCTCAT TCTTCTCAGT GCTGGCCACA GGTGGGGCTT GGCTGCTGGC  
 38551 TAAAAGGTGC CGGGGAGGGG GATACAGCAG CTCCCAGCTC ATCCTCAGAG  
 38601 GGTCTGGGA TCAAAGGTAT TTACACCCAG GGATATTTCA GATAAATCTT  
 38651 TTCATCTATG TGGAAAACAT ACAAAGTGGC GCAAGTGAGA AACTCCGATT  
 38701 TCCTAAGGTT GACAAGTCAA GTGCAGTAAT GATGTCATGG TAACCAATAT  
 38751 GTTTCCAAAC TTTCTAAGG TTGACTAGCC CCATGCACTT TGAGAAGTTG  
 38801 GTAAATAGGA TTGTCTGTCG TTTATAAAAT TGAAGACAG GTGTCTTGCA  
 38851 ATCACAGCCA CTCACAAAGG AAGCCAGAGA TGGTCCCAGC CCCTCCGCAG  
 38901 ACTTCTGTG GACTCAGGAC TGGTGGTCTC TCCTGGGCTT TGCTGTACCC  
 38951 GGCAATCCA GGGGCACAGA CTCAGGGTTC TGCCCTGCCG ACAGATGCTG  
 39001 CCTAGCCTTC TGTGTGTCAT AAGTCAACTC CCGCTCAGCC CCAGGCTGCT  
 39051 GGGTCCCTGC TGTGGGCCAA AAACCAGCCA CTTCTGCTGGT TTCTATCCCC  
 39101 CACCCCGTTC CCGAGGGAGG GGCTCTGGTG TGAGACACCC CCTCAGAGAG  
 39151 GAAAGTGTCT CCCAGCTTTG GAGAGAATCG AGGTGTCTCT TCTCTCTCTC  
 39201 CAGCTGGAAC CAGTGGGCCG CAAGACTACC TCCAACAAC TCACCCACAG  
 39251 CCGAGCCAAA CTCAAGTGCC CCTCTCTGT GAGTAAACGT CCTGCCTGCC  
 39301 CCAGGTGGAA CAGATGCCGT GGGTGGGGGT TGTCTGTCTC CCTGGAAGCA  
 39351 CAGAGGAGTC CCCGGGGATG CTCCCTCAA TGGCGCTTCA CTCACTGCCG  
 39401 TCTTCTCAA TCCCACCTGT CCCCAGTGCC ACGGAAACTT CTCAGTGTGT  
 39451 GGGCAGCCAT GGAGGGAGGG GAGAGGACGT TCAACAGCTC CAACCGAAGG  
 39501 GAGGACAGTC GCTCAGGGAG GCAGATGAGC ACTGGCGGGT GTCTCGGGTC  
 39551 ACCCATTCCT TCCGAAAGCT CTGATGCATC CTCAGTCTTA AAAGTGCAAC  
 39601 AAGGCCAGG CTGGTGGCTC ACGCTATAA TCCCAGCACT TTGGGAGGCC  
 39651 GAGGCGGTCA GATCACCTGA GGTCTGGAGT ACAAGACCAG CCTGGCCAAC  
 39701 ATAGTGAAAC CCCGTCTCTA GTAAAAATCC AAAAGTCAGC CGGGCATGGT  
 39751 GGGGGGCGCC TGTCTATCCA GCTACTCAGG AGGCTGAGGC AGGAGAATCA  
 39801 CTTGAACCCG GGAGGCAGAG GTTGCACTGA GCTGAGGTCA TGCCACTGCA  
 39851 CTCCAGACTG AGTGACAGAG TGAGACTGTC TCAAAAATAA AGTGCATCAA  
 39901 GCAGCTGTCC CGTGACAGGC AGTATACTAG GATCTGGGGA TCGGGAGGCA  
 39951 AAGATAAAAT AGACTCAGTG TCTGTTCTCT GAGCCTGCAA TGGTCTTCTT  
 40001 CCCTCGCCAC ACCCACTGCC CTTGCCTGGC CCACCTTCGA AGCCTGTGAC  
 40051 TTGTCTCCCC AGCTCTCTCT TCCCTCTTCT CCATCCACCC TACACTTGCT  
 40101 GCCAGACACA GATAGACCTT CCTGGAAATA ACTTGCCCCA TCAAGGCTGC  
 40151 TTGAAATCCT TGCTGTATCC CTAATGCCCA TTGACCAGAG TCTGGAGGGA  
 40201 GGGTCACTC CCTCCATGAT ACACACTGCA CTCCTGGCCG GTGGATCCAT  
 40251 CTCCCAGGAA GCCCCACGAC TGCCCGCATC CAGGCCTTTC CTTTTGCCAT  
 40301 CTGTTCTCTG AGGTTCATCT TCCATCTGCT ATGAGAACAT CCGCCTCCCT  
 40351 CCAGGTCCAG ATGTTGCCCT TACTAAGCGA TGGTTTCACC GTCTCTTACC  
 40401 TACCATTCTT GTCTCCAGAC ACTGACCCAT GTGGGTCTCC TTTTCTATTT  
 40451 GTACCTCTCA TGAGACACCG ACCCAGTCTC CTTTATGATG TGATTGTTTC  
 40501 TGCACATCTC AACTTCTCTC TGGGCCACAA GAAAAGATGT CACATCTTAA  
 40551 CCCTCCAGTC TCATCACAGC TTCCAGCAAG GGGGCTAAAC ACAGCACGTG  
 40601 CCCAATTCAC ATTCAGTGA AGGAGAGTGG AGAGGGGCAT AGGAAGGCAA  
 40651 GAACGCACAC GATCTGCCCA CATGCCTCCC CTCCCGGCC TTCTGATTTG  
 40701 GGGATCTTTC ATCTACTACA AAACAGCTG TCCTTCCATG CTGCCCTTCC  
 40751 CTGATTTCTG GGTAGTCTCT GGATGGGAGA ATGGGGACAG TTGTGACCAC  
 40801 GAGGAAGCAG AGGTGGGAGT TCTACAGGCC CCACAGGGCT CTCTGCCATT  
 40851 GGTCACTTAT CAGTTCCCAA TCTTTCAAAA TCAGGTTTGA TGGCCAAGGA  
 40901 AACGCTGGTG AGAAACCAA AGAAGGTTCT AGCTGGGTGT TGACCTCTTT

FIGURE 3, page 13 of 33

```

40951 AGAGGCCCAT CCCGCTAAAG AGGGTTTGGG CACAGCCTAA ATGAGGGAGC
41001 TTTACAAAAG GGAAGCTCTG TGAAAACGTG CAGGGTTATC GCAGCATCTC
41051 AGGAATGGGG ACTAGGCAAG TCTTGGCTTG GTGATGGATG GTTCACGGAG
41101 ATCCTTTTCCA CTGACCCCGG CTCTCTCTCC ACAGGAGAGC CCTTACCTCT
41151 ACACCCTGCG GAACAGCCGA TTGCTCCCAG ACCAGGCTGA AGAAGCCCCC
41201 GAGGTGCTCT ACTGGGCTGT CCCAGTGGCA GCGGGAGTCG GCCTTGTGGT
41251 GGGCATCATC GGGACAGTGG TCTGGAGGTG CAGGAGAGGT GGCCGGAGGG
41301 AAGATCCTCC AATGAGCCTG CGCACTGTGG CCTCTAGGC CCGGGGGTGG
41351 GTCCTCACCC TAAACTCCCT ATAGCCACTC TCTTCACCGC CCTCTGCCCC
41401 AGCCACTCCC GGCCACCAGG ACATGCTTCA ATGCCTGGTG CCATAGGAAG
41451 CCCAGGGGAC AGTCACAACT TCTTGGGGCC TGGGCTTCTT CCAGGCCTAT
41501 GTCCTGGAA TGATACATT TAAATAAAGT CCAAAGCTAT TTTATTCTTG
41551 GGTTTGCCTG CGTGAAGCAC TCACCTTCCA TCTCTGTGC AGCCAGGTG
41601 TGGGAGCTGC CACTTTTTGT GGCCTGCCTC CAGCAGGGCT GCCAAGCCA
41651 CGACCAACCA GAGCCCAAAC TGCCTGCCAC CACGAGCATA TCCTCAAGTC
41701 ACCAAACCCA CTATTTCAA GGCAGAAAAA ATGCTGGTCA CCAGGTGGTG
41751 GCTGGAATTT TGGAGCTGGC TGGTTGCCAT TCAGTCCAAT CCAACACATA
41801 CCTATTAAGC AACTGTTTTG TATCCAGGAC AATGCGAAGC ACTGAGGTGC
41851 CTCCTAGGCT GTGCATGTCG CAGCCTGGCA GAGAGGTCAA ACTCCTTCAA
41901 TAACCAAGAA GCCACGTGAT GATGTGTAAC TACTAGGGCA TCAGTAGGTA
41951 AATGTGTCTG ATTGTTTTAA AGAATAGAAA GGGTCTTCG GGGAAAGTTT
42001 CTGGGGGAG AGCAACCTTC ACATGTCATT TTGGGAAAAG GAATAAAAAA
42051 TGATTGGGAC ACAAATACCT CCTATATTCT CAACCTGATT TTCTCAAGGT
42101 GCTAAATTTA GGAAAAAATT CCTATTTCTA TATGCCCAGG TTTCTGAGGG
42151 AAAACTAGAG AGAGCTGAA AATATGGGCT GCATTCACTG AGCCCTGCT
42201 AGGGGCGAGG CCCCGTGTG GAGGCCTTCC ACAGATGGTC TCTTTTATGC
42251 TGCACAAAAG CCCAGGGAGG GGGTAAAAGG AAAATCTTTG AAAATAGAAG
42301 TGATGCTTGC GCAACACCGT GAATGTACTA AACGCCGCGA ATTGTTCAT
42351 TTAAATGAT TAATTGTGTA TCATGTGAAT TTCACTTCAA TAAAAAGAA
42401 TCCAGGGAGG TAGACATCAT CTGCATTGTA AACCTCTCTC TGATCCTGAA
42451 GTCCGGGATG ATAAAGAGCC TGAGTCACAA TCCCGGATGC AACACTGAAA
42501 TGCTGTGCCC TGAAGCTGCC TTCGCCAGCC TGAGCCCAGT GTCCCAGGCT
42551 CTGCATCTGT AAAAAGTGA GTAAGAGTAC ACATTTTGCT TATCTCACGG
42601 CGCTGCTGAA AAATAAGGAA CCGTGTGTGA ACCTCTAAT CTAAATGCT
42651 GCACAACTGA AAATGGCCTT TTTCTCGGT GAAGAGTTGG GATAAGGCCC
42701 AGACTGTTGG GGAAGATGTG AGACCCAGAG ATGAGTTTGG GGAATGGGG
42751 TAATAACATA TGGGTGGAGA GTGCCCGCCT TCCTCTCAGG GAGGTTTCATC
42801 ACCTTATCTC TTTCTGTCTC AACAGAGAAC CCGGAGGACC TATACCCAGT
42851 TCCGTGTTCT TCTGGGCTTC AGTGTCTGTT TCTATACAAT GGGAACAGCA
42901 TGCATTCCCC TGCTTTTTCC TATAGACTGG AAAACGTGGT GACCAAGTCA
42951 CACATCCCAG CTATGCTCC CGGCTTAAGA CAGTGTAAAG ACAAGGTAA
43001 CCCTTACACT CCTGGTTTGA GACAGTATAA CGACAAAGGT AACATAGGAA
43051 GTCAAGGAGT TCGCTTCACC GCCCTCCCC CCACCCACC CTTTTTTTTT
43101 CCTGCAAGTT TCTATTCTTC CCGCAGCTCC TACCTCAAAG CAGCATGGAT
43151 TCATAACCAC AGGCTCCCCT CATTAGGGCT TGGGGAGGGA GGGTGTGGA
43201 ATCCCACTG CCAGAGTAAT CCAGACTAAA ACATCAACAA ATGGTCCCAG
43251 CTGGTTCAAC AAGGAACACT TGGCAAAACA AAGAAATCCT GTCTGGAGCG
43301 ACACGGACAC AGCCACAAAC CAGTCACCAA ATTCCCAGCA AGTATGTGCT
43351 AAGAAGCCAA AAATTAAAAA TACGTGAGAA GCACCCACTT GAAATTGGTG
43401 GTATTACATA CATACACTGG CTGTGCGCCA TGGGGTTTTT CTGTAGGAAA
43451 ATGTCCAGTC TAGCTAGAAC GGCACCCACA GCCACCCAT GAGCAAAGCC
43501 ACCAAATACC TGAGGAAGCC ACAGTCCATG GCACTCCCA TGG

```

**FEATURES:**

```

Start: 3000
Exon: 3000-3098
Intron: 3099-4565
Exon: 4566-4637
Intron: 4638-12578
Exon: 12579-12644
Intron: 12645-16277
Exon: 16278-16338
Intron: 16339-16891

```

FIGURE 3, page 14 of 33

Exon: 16892-16950  
Intron: 16951-18968  
Exon: 18969-19036  
Intron: 19037-19303  
Exon: 19304-19344  
Intron: 19348-19582  
Exon: 19583-19682  
Intron: 19683-23297  
Exon: 23298-23349  
Intron: 23350-24668  
Exon: 24669-24749  
Intron: 24750-26803  
Exon: 26804-26869  
Intron: 26870-27334  
Exon: 27335-27426  
Intron: 27427-28826  
Exon: 28827-28905  
Intron: 28906-30276  
Exon: 30277-30381  
Intron: 30382-31185  
Exon: 31186-31281  
Intron: 31282-39203  
Exon: 39204-39278  
Intron: 39279-41134  
Exon: 41135-41335  
Stop: 41336

CHROMOSOME MAP POSITION:

ALLELIC VARIANTS (SNPs) :



21406	G	A	Intron			
22401	T	C	Intron			
22926	A	T	Intron			
23007	G	A	Intron			
23180	C	G	Intron			
26490	A	G	Intron			
26505	A	G	Intron			
29336	T	C	Intron			
29829	C	T	Intron			
29830	G	T	Intron			
29840	C	T	Intron			
29944	C	T	Intron			
30468	T	-	Intron			
30471	T	-	Intron			
30802	G	A	Intron			
30894	G	A	Intron			
30907	G	A	Intron			
31447	C	A	Intron			
31603	G	A	Intron			
31685	A	G	Intron			
31833	C	T	Intron			
31970	C	T	Intron			
32177	A	G	Intron			
33018	T	C	Intron			
33090	T	C	Intron			
33993	C	T	Intron			
34284	G	A	Intron			
34314	G	A	Intron			
35392	A	G	Intron			
35599	T	A G	Intron			
35997	C	A	Intron			
36085	C	T	Intron			
36270	C	T	Intron			
36481	G	A	Intron			
36619	G	A	Intron			
37088	G	C	Intron			
37204	G	A	Intron			
37485	G	A	Intron			
37624	-	A	Intron			
37685	C	T	Intron			
37769	C	T	Intron			
38897	G	A	Intron			
40155	A	G	Intron			
40355	G	C	Intron			
40486	T	C	Intron			
40512	A	C	Intron			
40622	-	A G	Intron			
40654	T	C	Intron			
40933	C	G	Intron			
41171	T	C	Exon	418	L	L
41379	T	C	Beyond ORF(3')			
41388	C	T	Beyond ORF(3')			
41880	A	C	Beyond ORF(3')			
42278	G	A	Beyond ORF(3')			
42339	G	A	Beyond ORF(3')			
42612	A	G	Beyond ORF(3')			
42817	T	G	Beyond ORF(3')			

Context:

DNA  
Position  
1573

TGGCTGTCCCCACACTGGAGATGCCCTCACCTCCTGGTCTGGCCCACATGCAGTGGTGAT  
GCCTCAGGGTCTTTGTGACTTGGTCTATCCATGTGTCCAAGTCTGTAAAGGAGGACTTCT  
GCCAGAACGTCCCCCTCCAGAGGCTGGAGCCATGACTCCCCCTGTACCCAACCTCAAGGT  
GCCTGGCAGGAACCTCTATGATACCAGGCAGCCACAGAGGGGAGGGATCAAAGTTGGGAC  
AGAGGCTGGTGTGTTGAGAGACAGGATAGCCTAGACTGTGAACATGGGCAGTGGTTAGGGA  
[A, T]  
GTAGACATATGTGGTCAAACCTGTAACAGAAAGCAAGGAAAAGGTACAAGCAACTCAGTTA  
CCTTTAGGGGAAGAAGAGAATTAGGAGGGACACAGGGAGCTTCAAACCTGGGAGTGTGTTG  
TTTCTTAAACTGGGCCATAAGTACATGGATGTGTGTTTTATTATTCTTTATATCTTACAC  
ATCTATTTACTCAGCAAATCTTACAGAACTTCCCTGTGTACCAGGCATTGTTTCAAGTGCT  
TTAGAAATCTCTCTCTTAAGTAGATGTGATGGGTGTGAAATAATTATGATGAAACCAA

1737

TACCCAACCTCAAGGTGCCTGGCAGGAACCTCTATGATACCAGGCAGCCACAGAGGGGAG  
GGATCAAAGTTGGGACAGAGGCTGGTGTGTTGAGAGACAGGATAGCCTAGACTGTGAACAT  
GGGCAGTGGTTAGGGATGTAGACATATGTGGTCAAACCTGTAACAGAAAGCAAGGAAAAGG  
TACAAGCAACTCAGTTACCTTTAGGGGAAGAAGAGAATTAGGAGGGACACAGGGAGCTTC  
AAACTGGGAGTGTGTTTGTCTTAACTGGGCCATAAGTACATGGATGTGTGTTTTATTA  
[T, G]  
TCTTTATATCTTACACATCTATTTACTCAGCAAATCTTACAGAACTTCCCTGTGTACCAGG  
CATTGTTTTCAAGTGCTTTAGAAATCTCTCTCTTAAGTAGATGTGATGGGTGTGAAATAAT  
TCATGATGAAACCAAAGGGGACACAGTAGGGCACTCATGTGAAAGAAGGAGAGGTCTAAG  
GCATAGCATCAGAGGCCCCAAATATCAGCTCCAACACCAGAGGATGCATTTTCTTTTAA  
ATTAACACTAAATTTTCACTGCCCAAATTCATTTGCTCAGCTGAATAATCGGTTGCAGG

2498

AGGGCGGGAGGGAGCCTCTCCACCCCTCCCCTGAACCTGGGCAATCAGAACCAGCCCCTG  
ATGGAAGCCTGAGCTCTGGGGCCTCCTGCCTCCCCCTCTTTGTGCAGCGTTTTGTGTAAC  
TGCGTTCTGACCCCTGCGGGAGAACTCCCAAGAGCTAGCCAGGCTGGAGGCCTTCAGCCGA  
GCCTACCGGGTAAGACCAAGAAGGGCACCATGCTGTGTCTCTCCCCCTACGTTCACTCTA  
ACACACAGCCCAGAGCCCCTAGAGGAGGCACACAGGGAAGGAAAAGCTGGTCAGGGATTG  
[T, C]  
GGGGAGACGGGGAGCAGCCTGGGTGCCTTCCCTCTGTCTCACGTGACTGTGGTGTCTCAGG  
TGCCCTGGTTGGAATCATCCAGTAGGATCCAGGTGGAAGCCCTCATGGCCAGCTAC  
CGTTGAGGGCTTAACCCCAACTCCTGGCCCGTAGCCCTGGATGCCTCATGAGACCACCTT  
TCCCTCCCCCACTCCCACTCCAAAGGCAGGTGCCGAGCCTCTGGAGGTTCTTCCAGGTT  
TTTATCCCTTTTGGGACTTCCCTGCCTAGCCCTTCAGAGAGAGTAGTCTACTTACAATCAA

3097

AAAACAAAAGGTGACCCAACTGTTTCCAAATTCTCTGGAAGGGACTTGCCCTCAGGT  
GATTTGTGTTCTCAAGGGAAAGGCTGAGTCGGCCCTCCATCCAGGGAGATGGACTGCCC  
ACCACCCCTACTCTTGCCCTCACTGGGTCTGGGCCACCCAGGGCCTGGGCTGAAGACCC  
TGTGCATGTGTCCCCAGAGCAGCATGCGCGAGCTGGTGGGGTCAAGCCGCTATGACACGC  
AGGAGGACTTCTCTGTGGTGTGTCAGCCCTTCTTCCAGAACATCCAGCTCCCTGTCTGG  
[C, T]  
GGTATGTCCCCTGCCCCTCGCCCATGGTACTCTTTTAGAGGAAGAAATGCAAGGCAGAATT  
GCCAGTTGCTTCCACGAGCATGTGCATAAAATGGGAAAGACACAGCTCTCCAGACGCTG

3116

ACCTGTTTCCAAATTCTCTGGAAGGGACTTGCCCTCAGGTGATTTGTGTTCTCAAGGGA  
AAGGCTGAGTCGGCCCTCCATCCAGGGAGATGGACTGCCACCACCCCTACTCTTGCCCT  
CACTGGGTCTGGGCCACCCAGGGCCTGGGCTGAAGACCCTGTGCATGTGTCCCCAGAG  
CAGCATGCGCGAGCTGGTGGGGTCAAGCCGCTATGACACGCAGGAGGACTTCTCTGTGGT  
GCTGCAGCCCTTCTTCCAGAACATCCAGCTCCCTGTCTGGTGGTATGTCCCCTGCCCTC  
[A, G]  
CCCATGGTACTCTTTTAGAGGAAGAAATGCAAGGCAGAATTGCCAGTTGCTTCCACGAGC  
ATGTGCATAAAATGGGAAAGACACAGCTCTCCAGACGCTG

4823

AACAACTACTACACCCGTGTCTCTCTTTCTTCCCTGATCAGCTTGAACCACTTGGAAG  
CAAAACAGAGACCTGGACCTGAGAGCAGAGATGCCATCACCTGTCCCACTCAGGTAGT  
AGGGGAGGACCTGCCTGGCTCTCTCCACAAACCAGGGCACACAGCTCGCCCTACCCACT  
TCGTCTCCACCACAGCTTCTCAGTACCCATCTTGCCCTTACTGAGGCCTGAGAGAT  
TTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGGTGAGAGGGGAGACAAGAGT

FIGURE 3, page 17 of 33

[G, A]  
CAGCTGTTCATTGGGAACAGGAGATGCAGCAGGGAGAGGAGCCTGGGCCCCAGCAGAGGG  
AGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAAATCAGGATGCCAGGAAAATG  
GCAGGAGGGCTTCTCTAGCAGTGGTGTGGGGCAGATGAAAAATCTGACTGCAGGTT  
AGAGGGCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGGTTGGAGAGAGGAGAGCCTG  
GATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAAAGGGCAGAGCCCCG

4924 CCTGTCCCACTCAGGTAGTAGGGGAGGACCTGCCTGGCTCCTCTCCACAAACCAGGGCAC  
ACAGCTCGCCCTACCCACTTCGTCTCCACCACAGCTTCCTCAGTACCCATCTTGCCCC  
TTACTGAGGCCTGAGAGATTTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGG  
TGAGAGGGGAGACAAGAGTGCAGCTGTATTGGGAACAGGAGATGCAGCAGGGAGAGGAG  
GCCTGGGCCCCAGCAGAGGGAGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAA  
[A, G]  
TCAGGATGCCAGGAAAATGGCAGGAGGGCTTCTCTTAGCAGTGGTGTGGGGCAGATGA  
AAAAATCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGG  
TTGGAGAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGC  
AGCAAAGGGCAGAGCCAGAGCTGTATGGAGGAGGGACGAGGTTGGGCTTACCAGGACAC  
GGCAGCTCCAGGCTCCTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGCTTGACTTCAGG

4989 TCGCCCTACCCACTTCGTCTCCACCACAGCTTCCTCAGTACCCATCTTGCCCCCTTACT  
GAGGCCTGAGAGATTTGGAGGATGGAGGGGAGTCCATGAGGATGGACAGGGGAGGTGAGA  
GGGGAGACAAGAGTGCAGCTGTATTGGGAACAGGAGATGCAGCAGGGAGAGGAGGCCTG  
GGCCCCAGCAGAGGGAGAGGATCCCGGTGAGAAAAGTGGGCTCCTGAGAGAGGAAATCAG  
GATGCCAGGAAAATGGCAGGAGGGCTTCTCTTAGCAGTGGTGTGGGGCAGATGAAAAA  
[A, G]  
TCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCAGGCTTAAGAGCTGTGGTTGGA  
GAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCACAGAGGACAGCAGCAA  
AGGGCAGAGCCAGAGCTGTATGGAGGAGGGACGAGGTTGGGCTTACCAGGACACGGCAG  
CTCCAGGCTCCTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGCTTGACTTCAGGCCTGG  
GGTGGGGCAGGTTCTCATTGTCTTCAGCTCCTGTTCTAGGCCCGGTCTTATGGCTTTT

5274 GGGGCAGATGAAAAATCTGACTGCAGGTTAGAGGGCCAGGCAGGAGCCAGGCAGGCTT  
AAGAGCTGTGCTTGGAGAGAGGAGAGCCTGGATTAGGGAGATTCCACAAGGAAAGGATCA  
CAGAGGACAGCAGCAAAGGGCAGAGCCAGAGCTGTATGGAGGAGGGACGAGGGTGGGCC  
TACCAGGACACGGCAGCTCCAGGCTCCTTTAAGGAGGAATCCGTAAGTGGTTGTTAAGC  
TTGACTTCAGGCCTGGGGTGGGGCAGGTTCTCATTGTCTTCAGCTCCTGTTTCTAGGCC  
[C, T]  
GGTCTTATGGCTTTTTAACCATAAGGCCAAGGCCAGAAAACCTCAGCAGCAATAAAA  
GCAGAAGGCCCTGACCAATCTGGGAGGCTGGGTTTCCCTCCTAGGTGGGCCACACCACCC  
TCTCCACCCCTCCCTGCTGGGGAATGGACCTGCAGCTCCCCATGTGTCTGCTGGGAATC  
CTGAGAGAGTGGGCACCCCTGTTACATGCCTGCTCCCTGTCTGCTGCCTGCCCTACCCC  
AGTCTTGGGCTCAGGCTCAGTCTTGTGTGCCATCAGCCCCATCAGGAGAGCAAGAATGGC

5792 CTGTCTGCTGCCTGCCCTACCCAGTCTTGGGCTCAGGCTCAGTCTTGTGTGCCATCAGC  
CCCATCAGGAGAGCAAGAATGGCAGGAAGAAGGGATGGGAAGTGAAGACAGTCGTAGCAG  
AGGGCTCAGTTGCTGGGTCTTGTGCTTGGAGCTAAGGAGATTGTGAGATTCTGCAACAGC  
TAGTGCAACACAGATGCCTTAGTCCAGGTGGTCAGGTGCTGGCCAAAGGCCTGGAGCAA  
AACCTTAGAGGGCCCTACTGTGCCAGGTGTAACCTCTTAACTGCTTTCCTAAGGATGCC  
[-, T]  
TGGGGGTTCTAGGGGAGCAGCCAGGGACCGTGGATAGTGGGGGCATTGGGGACTCAGAA  
ATAGCCATATTGTAGATATTTCAATATTTTACCAACCTATAGCCATACTGAATATCAGC  
CATGGAGGGCCCTTTCAAACCTGTCCACTCCCCTTCCATTACATAACAAAAGCAGCCATC  
ATTTGCTCTTTCTTTCAACAAACGTGTATTGAGTACTGAGTTGGAGCCTAAGCACTGGGT  
CAGGGAGAGCCCTGTCAACCTGGGCTTCGAGGCAACCACTTCAGGCTTTACCCAGATC

5871 TGGCAGGAAGAAGGGATGGGAAGTGAAGACAGTCGTAGCAGAGGGCTCAGTTGCTGGGTC  
TTGTGCTTGGAGCTAAGGAGATTGTGAGATTCTGCAACAGCTAGTGCAACACAGATGCCT  
CTAGTCCAGGTGGTCAGGTGCTGGCCAAAGGCCTGGAGCAAAACCTTAGAGGCCCTACT  
GTGCCAGGTGTAACTCTTTAACTGCTTTCCCTAAGGATGCCTTGGGGGTCTAGGGGAGC  
AGCCAGGGACCGTGGATAGTGGGGGCATTGGGGACTCAGAAATAGCCATATTGTAGATA  
[G, T]  
TTCAATATTTTACCAACCTATAGCCATACTGAATATCAGCCATGGAGGGCCCTTTCCAA  
ACTGTCCACTCCCCTTCCATTACATAACAAAAGCAGCCATCATTGTCTCTTTCTTCAAC

FIGURE 3, page 18 of 33

AAACGTGTATTGAGTACTGAGTTGGAGCCTAAGCACTGGGTCAGGGAGAGCCCTGTCACC  
CTGGGCTTCGAGGCAACCACTTCCAGGCTTTACCCCAGATCAGGCAGAGACCCCCAAAAG  
GAGGCTGCTCCACCCAGCAGCATCTTAAGCTGAGTGGGCTCAGTGCCTCCCTTCTAGACA

6562 CTAAGCAACATTGGAGCCCATCTGAAAGGGTCCATCTGTTGGCCAGCCCACTTCACTG  
TGTTCTGAGCATTTGCAATTCCTCAGTCCCCTGCTCCCTCCCCTGTCCTTGGAGTGA  
TATAAAAGTCCACCAGCATCTCAGTGTGAGCTGACAGGGGCCAGGCAGCACCTATTTTGT  
TCCTAGATGTGTCTAAACATAGAGGCAACAGGCAACAGGCAAGACGCAGTGGGGGGCGGG  
AGGCAGGAGGCCGAGATGGCTGTGAGCATGAGCTTTCTCAGCCTCCTCCCTTCTCCCATC  
[C, T]  
GCAGTCTAACTGCTCATACGTTCTGTGTGCCAGGTAGGGTGACTTAACAGCACGCCATGG  
ATTTCTGTTGTAGTTTCAAGTTGGACAAATCTTTTACAGACAACCTTTTGACTAGCCTTC  
TGTGGACTGAGCCTATACTCTGCCTTAATGGGCTCTCTGCCCACTCCTTTCTAACCCTA  
GGGCAGCTGGCTGAACACCTGGTCTTTTCTTAGGTTTTCATTCTTTTGGACCTCTCTGAA  
GCCCTTGTCAAAGTCACCACCTCCCCCTTGAATTCACCTCCTTGGGTTTGTGGACA

9859 CAATAAACCCAGCTAAAAACAAGCCCAATAAAACCCCAATAAAACCCATTAGACAGGAACA  
TAGGAGTTGGAAAAAAGAAAAGAGGGGAGGGGGAAGAAAGCCCTGAGGCACCCCGGC  
TGCCTGTCTGCCACAACCCCTGGGCTGTAATTGTTCTTGGCATGGCCTCAGTCTGCAACAC  
ATTCTAGTGTCTCCTTGACCTCTAGCCCTCTAGCTCTGCCTCCCTTTCCCAACCTGTAG  
ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGC  
[G, A]  
CTCGCAGCACTTCCAAAAGGTCAAGTTTGTCTTCCCTCAGTGCCTCCCATTCTGGTCA  
CGGTAGGACTGACTCCAGCCCTGGACCTAAGCTGAGTCTGGGCTCCTTTGACGTGCAG  
GGAGAATGCCACTGAGTCTTGTCTCTGAGGACCTACCTCTCCAAATCTTGCCTCAGTTC  
CTCAGCAGGTACTACACTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTTC  
TCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTTCTCTCTC

9875 CAATAAACCCAGCTAAAAACAAGCCCAATAAAACCCCAATAAAACCCATTAGACAGGAACA  
TAGGAGTTGGAAAAAAGAAAAGAGGGGAGGGGGAAGAAAGCCCTGAGGCACCCCGGC  
TGCCTGTCTGCCACAACCCCTGGGCTGTAATTGTTCTTGGCATGGCCTCAGTCTGCAACAC  
ATTCTAGTGTCTCCTTGACCTCTAGCCCTCTAGCTCTGCCTCCCTTTCCCAACCTGTAG  
ATCTTGTGATCAAATAGATTCAATGAAACACATTGTCCAGTTGCACTCGCAGCACTTCCA  
[A, G]  
AAAGGTCAAGTTTGTCTTCCCTCAGTGCCTCCCATTCTGGTCACGGTAGGACTGACTCC  
AGCCCCCTGGACCTAAGCTGAGTCTGGGCTCCTTTGACGTGCAGGGAGAATGCCACTGAG  
TCTTGTCTCTGAGGACCTACCTCTCCAAATCTTGCCTCAGTTCTCAGCAGGTACTACA  
CTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTTCTCAAGTCTGTCTCCCC  
ACCTGAGCCAATTGTGAGTTTCTCTCTCTCCTCCTCTCATCCTGGCACCTAGAAATGCTC

10279 GGGAGAATGCCACTGAGTCTTGTCTCTGAGGACCTACCTCTCCAAATCTTGCCTCAGTT  
CCTCAGCAGGTACTACACTGACTGGCCATGCCATTCTCTGATGCTTCACTGCCTCAGCTT  
CTCAAGTCTGTCTCCCCACCTGAGCCAATTGTGAGTTTCTCTCTCTCCTCCTCTCATCCT  
GGCACCTAGAAATGCTCTTAACGCTTGAAGTCTCAACCAGCATGGGTCACTTGTATTAT  
AGCATGCTCCAGATCGCCCTCTTTGTTGGTGAATGCTCAGGGAATGCTTACTGTTAACC  
[C, T]  
GAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAAGCCCTCTCCTGTCTTATGCTG  
TTGTAGAGGGTCCAGGGCTCACTTCTCCCACTTGGCCCTGAGTACCTCTCCTTGAAAGGA  
TGTCAGGGGCTGGGCGCAGTGGCTCACGTCTGTAACCCAGCACTTTGGGAGGCTGAGGC  
GGGCGGATCACCAGGTGAGGAGATCGAGACCATCCTGGCTAACATGCTGAACCCCCGTC  
TCTACTAAAAATACAAAAATAAAATAGCCATTGTGGTGGCAGGTGCCTGTAGTCCCA

10474 TCTCTAACGCTTGAGCTGCTCAACCAGCATGGGTCACTTGTATTATAGCATGCTCCCAGAT  
CGCCCTCTTTGTTGGTGAATGCTCAGGGAATGCTTACTGTTAACCCGAGACAAGCCCAAG  
TAGCTACATGGACCTGCCACCATAAGCCCTCTCCTGTCTTATGCTGTGTAGAGGGTCCA  
GGGCTCACTTCTCCCACTTGGCCCTGAGTACCTCTCCTTGAAAGGATGTCAGGGGCTGGG  
CGCAGTGGCTCACGTCTGTAACCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACCAG  
[-, G]  
TCAGGAGATCGAGACCATCCTGGCTAACATGGTGAACCCCCGCTCTCTACTAAAAATACA  
AAAAATAAAATAGCCATTTGTGGTGGCAGGTGCCTGTAGTCCCAGCTACTCGGGAGGCT  
GAGGCAGGAAATGGCATGAACCCAGAAGGCAGAGCTTGCAGTGAGCCGAGATCGCGCCA  
CTGCACTCCAGCTGGGCAACAGAGCAAGACTCCGTCTCAAAAAAGCAAGCAAGAAAGAA  
AGGATATCGGTTACCTGTTTTCAGACAGGAATGCTGAGACCAGGGAAGGGGAGACTTGTC

FIGURE 3, page 19 of 33

10566 CTTACTGTAAACCCGAGACAAGCCCAAGTAGCTACATGGACCTGCCACCATAAGCCCTCT  
CCTGTCTTATGCTGTTGTAGAGGTCCAGGGCTCACTTCTCCCACTTGGCCCTGAGTACC  
TCTCCTTGAAAGGATGTGAGGGGCTGGGCGCAGTGGCTCACGTCTGTAACCCAGCACTT  
TGGGAGGCTGAGGCGGGCGGATCACCAGGTGAGGAGATCGAGACCATCCTGGCTAACATG  
GTGAACCCCCCTCTCTACTAAAAATACAAAAATAAAAAATAGCCATTTGTGGTGGCAGG  
[C,T]  
GCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAAAATGGCATGAACCCAGAAGGCA  
GAGCTTGCACTGAGCCGAGATCGCGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACT  
CCGTCTCAAAAAAGCAAGCAAGAAAGAAAGGATATCGGTTACCTGTTTCAGACAGGAATG  
CTGAGACCAGGGAAGGGGAGACTTGTGCGGTGCCTCAGGGAACAGTATCTGAGCTGGG  
GGCTGAGAGCTCTGTGTGGGTGGACTCTGTCTCCAGTCGCTGCTGAGTCCCTCTCTTC

12038 CTGGGAATACAGGCATGGGCCACTGTGCCAGCCTGGTTTTTTCTTCTTGTTCCTATTTTA  
TTCTCACATTTTTCAGACCATGGGCTTACTACTCCACTGAGCACATTTTGTGAGAGTGCTC  
ACAGCCCTGGGCCCCGTTGCTGTTTCTCTGATCTCAGTCTTATCAACTTGATCTTGCTTTG  
CTGTCTCAAAAATATACATTTTCTCATTAGCTTTCTCCCACTTTCTCTTTGTCTGCTTCTTC  
TTCCTTCTTTAACTAACTCCTCACCTGCAACTGGGGGGACTTGGATTCTTGACTGGGCTT  
[G,A]  
TGTGAAAACCTGATTGTAAAACAGATAGGTAAGTAGGGAATGAGGAGGGTGTTTTACAAGA  
AAAAAAAATGACTAAGATACAGGAACCCCAACCTAAAGAGGAAAAGACATACAGTTCAAA  
GGAGGCAGAAAGAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTTCTG  
GAAGATTAAAAAATGCTGAAACATGAATCCCTTGCTAGAGAAATTACAAAGCCCAAGAAA  
ATAGATAGGTCTGAGGATTAGGGAGCTGTTTCACTTGCTAGGAGGAACACAAAAGCACAGA

12159 CAGCCCTGGGCCCCGTTGCTGTTTCTCTGATCTCAGTCTTATCAACTTGATCTTGCTTTGC  
TGTCAATTATACATTTTCTCATTAGCTTTCTCCCACTTTCTTCTTTGTCTGCTTCTCTTCT  
TCCTTCTTTAACTAACTCCTCACCTGCAACTGGGGGGACTTGGATTCTTGACTGGGCTTG  
TGTGAAAACCTGATTGTAAAACAGATAGGTAAGTAGGGAATGAGGAGGGTGTTTTACAAGA  
AAAAAAAATGACTAAGATACAGGAACCCCAACCTAAAGAGGAAAAGACATACAGTTCAAA  
[G,A]  
GAGGCAGAAAGAAAACATTACAGATACTCAAATATATTGATAATCATAACACTTTCTGG  
AAGATTAAAAAATGCTGAAACATGAATCCCTTGCTAGAGAAATTACAAAGCCCAAGAAA  
TAGATAGGTCTGAGGATTAGGGAGCTGTTTCACTTGCTAGGAGGAACACAAAAGCACAGAC  
CCCAGACTACAATGGGTATGAAACCTCTGCACGCCTTTTGTGTCCATCCCTTGCCAAA  
GCTGTTATGTAAAACCTCCGGGGGAATGAATGAAATTATGTTTATACAGTTCTTTCTAT

12979 CAAACCATTTCACCTGCCAGGGGCTCGGGTGTGGTACAGGTTTTCAGAGTATTCAGTGAA  
GCAGAAATGTACTTCTTACATACTGGGGATTGGAATGTACAGAAAAGGCTCCCGGACCAC  
GAAGCCCCAGGATTGTCTAACATGTTCTCAAGTTGCTTACCTGACGTGAGCCCCCAAGC  
AGAGGAAGTGTCTATGGATCGATTTTCTTTGACCTTGGCAATCCTGGGCTCAGAGACGTG  
GTTACTGCTTAGGCAGCTCAGCCTCTCAAGAGGGAGAGGCAGCTGGTGTGATGTGGCGTT  
[A,G]  
ACTTCTTGGAAGGTGGAGGCTGAGTGGGAGGGAACATAATTCTGGGGATGGGACCCAAA  
AGGAAGTGGAGGCACGTTGTTTCTGTTGCTGTTGGGCCCCCTAGGCCTGTTTGGTTCAAGT  
CAATCATTTCTAGTGTGAGGATTGAGAGCCCATGGTTAATTCCATTGGATTAAACCATGTC  
TGTGAGCCTAGGACGGCCACTGCAAGACGGCCTGGAGGACCCCGGACTATACCATGACT  
GGCAGTCAGGCCTGGTCCGGATCAGGTCTGTTGGTCACCAGGATGGGGTTTGACCCGACG

13865 GGGCTGGGGCAAGGGAGGCAAAATTCAGGTGCACCATGCAGGCCAGGCCTTCTGAGGTGA  
GATTTAACTGAGACATGCATAATGAGGAGACACTTGTATACAGGGAGCCAGGAACACA  
GTCCAGGCAGAAAGGACCATGGACCACACAGGCTCAGAAGTGGGACTGTGTTGGGTGTAT  
TTGGGGAAGAGAAAGAAGGTGAGAGTGGCTGGGGGCATGAGAATGAGGTGGAGAGTGGGG  
GAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAAAGCCTGAATTACTGAGTAAAA  
[C,A]  
CACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGGCGGTCTCAGGAGAGAGTGACAT  
GATCTGGTTTACGTCCTTTCAAAGATCTCCCTGACTGCTATGTGTAGAATGGGTTGGCCAT  
CAGCAGGAGTATTGGGGAAAGACATTTTATAAGCCAGCTGAAGAACTAACCATATGA  
AATCATTAAGAACTATTGGATGCTAAGCTCTGGGGTGCAGCAATACCAGATTGCTGGCT  
GCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAGGCTCACGTTAGCCAGTGGAG

13964 ATACAGGGAGCCAGGAACACAGTCCCAGGCAGAAGGACCATGGACCACACAGGCTCAGAA  
GTGGGACTGTGTTGGGTGTATTGGGGGAAGAGAAAGAGTCAAGTGGCTGGGGGCATG

FIGURE 3, page 20 of 33

AGAATGAGGTGGAGAGTGGGGGAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAA  
AGCCTGAATTACTGAGTAAAACCACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGG  
CGGTCTCAGGAGAGAGTGACATGATCTGGTTCACGTCTTTCAAAGATCTCCCTGACTGCT  
[A, G]  
TGTGTAGAATGGGTTGGCCATCAGCAGGAGTGATTGGGGAAAGACATTTTATAAGCCAGC  
TGAAGAACTAACCACATATGAAATCATTAAAGAACTATTGGATGCTAAGCTCTGGGGTGCA  
AGCAATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAG  
GCTCACGTTAGCCCAGTGGAGGCTTGTCTCATTGAACCACTGACCAAATCCCTGAGAA  
TTGAAACGTGAGCTGCATCTTGTGAATCAGGCATTCTTTCATTATTTCATTACCTATTG

14087 ATGAGGTGGAGAGTGGGGGAAATGAGATCAGGAGTGCCAAGGAGCCAGATCACACAAAGC  
CTGAATTACTGAGTAAAACCACTGGATTTCAAGTGGAGAAAGATGGGAAGGCATTGGCGG  
TCTCAGGAGAGAGTGACATGATCTGGTTCACGTCTTTCAAAGATCTCCCTGACTGCTATG  
TGTAGAATGGGTTGGCCATCAGCAGGAGTGATTGGGGAAAGACATTTTATAAGCCAGCTG  
AAGAACTAACCACATATGAAATCATTAAAGAACTATTGGATGCTAAGCTCTGGGGTGCAAG  
[C, A]  
AATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCTCTCTGAATTTTCTCAGGCT  
CACGTTAGCCCAGTGGAGGCTTGTCTCATTGAACCACTGACCAAATCCCTGAGAATTG  
AAACGTGAGCTGCATCTTGTGAATCAGGCATTTCTTCATTTATTTCATTACCTATTGGAT  
GCCTATGTAGAGTGGGCACTGCACTAAGTGTCTCGGTAGACAGTGGTGAGCCGAATGGGTC  
TGGATCTGCCCTCTTGGTTCTTCAGTCTCATGCATCTTGTCTTTGCTGTCTGGAAGAGCT

14309 CATTTTATAAGCCAGCTGAAGAACTAACCACATATGAAATCATTAAAGAACTATTGGATGC  
TAAGCTCTGGGGTGCAAGCAATACCAGATTGCTGGCTGCGGGTTATGCTGTGTCCAGCCT  
CTCTGAATTTTCTCAGGCTCACGTTAGCCCAGTGGAGGCTTGTCTCATTGAACCACTGA  
CCAAATCCCTGAGAATTGAAACGTGAGCTGCATCTTGTGAATCAGGCATTTCTTCATTT  
ATTTCATTTACCTATTGGATGCCTATGTAGAGTGGGCACTGCACTAAGTGTCTCGGTAGACA  
[G, A]  
TGGTGAGCCGAATGGGTCTGGATCTGCCCTCTTGGTTCTTCAGTCTCATGCATCTTTGCT  
TTTGTCTGCTGGAAGAGCTAAAAATCCAGAGCTAGAAGGGCGTGTGTTTGTTTTAAACAGC  
TTTCTACTCAAAGTAACCACAGAAACAAAATTCTGTCTGAGGTAACGTGAATGAGCC  
TAGAGGACATTACGTTAAGTGAATAAGTCAAGCACAGAAAGACAATACTACATGTTCT  
CACCATATGCCGAAGCTTAAGAAGTTGACTTCACAGAAGTAGAGTATAAATAGTGGTTAT

16028 ATATGCAAAAAGACATACAGGTATCCAGAAAAGACAGGCAGAAACCAGGAGCTTTACAAT  
TTTAAAATATTTGTGTTATTATTCTAAAAATATTTTAAATTATTGTCTAGGTTCTACCAT  
TATAATTAGTGTGCTAGTTAGCTTAATTTTATAAAACACACATACCTGTAATCTCATGTTAG  
GCATCCAAATGCTGTGTTCTTTGGGAGACCCACCTGTGTAGGACTTCATGTTTTCTTC  
CCTGCTTTGGGGCAGCCACTGGCTCCATTCAAAGCATAGATATATGGGGATAAGAAAGGT  
[G, T]  
GTGTGTGGGTGCACATGTGGAGACATGCACTATGGGTTGTGCATAGGGGTAGCTAGACAC  
ACCCATTTCTCCCCCTTTAATTTCCCTCCTAGCCACCTATAACTCACAGTTCTTTCCCT  
CACATGATCCTGTATGGTGACTCATTTCTAGCCTCCATCAAAAATCCCTTAGCTGGTTCT  
TCTTGGGCTGAAGCTTATCTCCCTGCACAATGAGTGTGGGCACTGAATCTTTTCTCCTG  
TTGATTTAGAACTGGGGCAGTGACTTCCTGTGTACAGAGTGGAAGGCTTCCAATAGTGT

16375 GGGTAGCTAGACACACCCATTTCTCCCCCTTTAATTTCCCTCCTAGCCACCTATAACTC  
ACAGTTCTTTCCCTCACATGATCCTGTATGGTGACTCATTCTAGCCTCCATCAAAAATC  
CCTTAGCTGGTTCTTCTTGGGCTGAAGCTTATCTCCCTGCACAATGAGTGTGGGCACTG  
AATCTTTTCTCCTGTTGATTAGAAGTGGGGCAGTGACTTCCTGTGTACAGAGTGGAAGG  
CTTCCAATAGTGTCCAACCTCTGGTGAGTGAAAACATCATCATCTCCTTCAATTAAGGG  
[C, T]  
CTTGCCGAATATCAGGTTGTGGGGAGACCTGCAAACATACCCTGGAGCTTTAAGCAGGA  
CTTGCTAATTTCCCTGCAGTGACAGCCTAGATCCTGCGGCTGCCGCCACAGCTGGGCTT  
CCATGTGGAGGTGCACAGAGCTCTCCATTGGATGCTACTTCTGTCTCCTTATAGTCCCA  
GTGGCAGTCCCTTAGGCCTCCCTGAGGAGTGGGAGGAGTACAGTCAAGGATTGGGATCTA  
CCTGCTGTGCTACATGACCTGCAGCTGGAACCTTCTGGACCACCCCAATGTCAATCA

16705 CTGCAACATACCCTGGAGCTTTAAGCAGGACTTGCTAATTTCCCTGCAGTGACAGCCTA  
GATCCTGCGGCTGCCGCCACAGCTGGGCTTCCATGTGGAGGTGCACAGAGCTCTCCATT  
GGATGCTACTTCTGTCTCCTTATAGTCCCAGTGGCAGTCCCTTAGGCCTCCCTGCCCA  
TGAGGCAGGTAGAGTCAAGGATTGGGATCTACCTGCCTGTGCTACATGACCCTGCAGCTG  
GAACCTTCTGGACCACCCCAATGTCAATCAGGCTCTTCTGAGGGTGGATGATAGCCATG

FIGURE 3, page 21 of 33

[G, A]  
AACCCATTCCCTGCAGTGCCTTGGTTGGTCTGAATGAATGGGAGGGGCAAACTGCTAAA  
GCCTTAAGCTGAAAAAAGTACAATGGGGAGCAGTGGGACAGAGTTATAGACTTCTGGTA  
AAATGTGTACTTTAAGAGGTAGATACCCCGAGCCCCACAACCACTCTCTGCTTGTCTC  
CCCTAGTCCACCAGCTCCGACCAGCAGACATCAAAGTGGTGGCCGCCCTGGGTGACTCTC  
TGACTGTGAGTAGTGAGCCATGAACCAGGATGGGCAGCTCAGAGTCCAGCCAGGCCCTGC

19708 CCTGCTGGAGGAGGGGAAGAGGAGGTTATCTGCAAGAAGGGAAGTCAGCCAGCCCTGAAA  
AGCCCCAGACTTCCCTGTGTCCCACCCATGTCCCCACCCTGCATGCTCATCTCAGTTACTG  
TGAGGGTCTGCAGGCTCTCACCTGTGCTCTTCTCCTCCTCCTCCTCTAAAGACATT  
CTGAAGAAGTTCAACCTTACCTCCTTGGCTTCTCTACCAGCACCTGGGAGGGGACAGCA  
GGACTAAATGTGGCAGCGGAAGGGGCCAGAGCTAGGTGAGTAGATGCCGTACAGGAGGGC  
[G, C]  
AG

21406 AAGAATGAATCTCCATCTCCAAAAATAATAAATTAATAAATAAAAGATACAAAGG  
AATCAAAAGATGAACCTCCCTGGCCACGAAGAGCTTGCACTCTAGGTAAAGGAGGCTAAACA  
AATGGGAATAACTTTCTGAAAAAGACAATGCTGGGTATGGCAACAATGCAGTGCTTCGC  
ATGGAGTACAATTAAGAGAACAGAAGAGCACACAGTATGAAGTGCAGTGTCTAAAGACAG  
ATGCAGACCCAGAAGGGACCCCTGAAATCATCCAGTCCAACCTCTCCTTTAAAGATGG  
[G, A]  
AAAAGTCAATCCTAGCAAGATTTCAGCAACTTGTACAAGCTCAACAGCAAGTTGGTAGCAG  
AGCTGAAAGTAGAACCCTGGTCCCTGGGGTAAAAAGGAAATGCAAGATGTGTGGATCA  
GGGAGCCCCAGAGAGGAGGCTCAAGGGAAAAGTAGGACTTGGTCTGGGCCTGAAGGATGGGA  
AGAAGATGGCTAGGAAGAGGGGAAGAAGCGGCATTTGTAAGTCCCTCCTACCCACGAG  
GGCTTATTGCCCATGGATTCTCTTAGTCACACCTTGAACCTGTTAAAGGTTAAAGGCAC

22401 AGGCCAAGGTAGGCAGATTGCCAGGAGTTGGAGACCAGTCTGAGCAACTTGGTGAAACC  
CTGTCTCTACCAAAAAATACAAAAATTAGCCGGGCATGGTGGCATGGACCTGTAGTCCCA  
GCTACTCGGAAGGCTGAGATGGGAGGATCACTTGAGCTCAGAGGGGTTGAGGCTGCAGTG  
AGCCGTGATCAGACCACTACACTCCAGCCTGGGCAACAGAGCAAGACCCTGTCTCAAAAA  
GGATACAATTTAATATTGTACCTGTGAAATCATCACCACAATCAAGATGAAAAATGTGTT  
[T, C]  
ATCACCACAGGAGTTTTCTCAGGCCCCCTTGGTAATCTCTCCCTCCTGCTCCTTCTGTC  
CCTACCTCACACCCAGGCAACCACTAACCTTCTTTCCATCACAATAGATTAGTTTGCAT  
TTTTAAAAATTTTATATAAATGGGATCAAAGAGTATATACTTTTTATCTGACTTATTAG  
CAAAATGATTTTGCATGATCCATGTTATTTCGGTATACCAATAGTTTCGTCCCTTTTAT  
GGCTGAGTGTAGTGTTCGGTGGCATTTCATATCGCTCATCCAGAACACCAATGGTATTG

22926 GTTCGTCCCTTTTTATGGCTGAGTGTAGTGTTCGGTGGCATTTCATATCGCTCATCCAGA  
ACACCAATGGTATTGTTTTATTTATGGCAGACATCAGGGGATGAAGGGAGAACTAATCC  
TGTCATCTCGTGTATTGGAGAGGGAGAAAAAAAAGTGAAGAGATGGGGAATGGTGC  
GGAAATCTAAGTAACACAGAAAAAAGAAAAACAAAGGATTAAAGGAGCAGAGAGCAGGGC  
TTAGAAGTAAAGTTAAAGGAGTCATTAAGCCTGGAAGGAGAAAACTGAGGGATAATTG  
[A, T]  
GAGCTGTGACTTTTCTCAAATATACAAAAGGTTATTTTTAAAACAGGCAACTGAAGAAGA  
AATGAACAGGCTTGGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAG  
GGGAGGCTTGACGGGATCCCAACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTA  
GATCAATGTGGCTCCAGGCTCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTCTT  
GAGAATGGAATGACCTTTGTACTGGTAACATCATTCTTCTGAAACACCTCTTTCCTAG

23007 TTTATGGCAGACATCAGGGGATGAAGGGAGAACTAATCCTGTCCATCCTGTTTATTGGA  
GAGGGAGAAAAAAAAGTGAAGGAGATGGGGAATGGTGCAGAAATCTAAGTAACACAGA  
AAAGAAAAACAAAGGATTAAAGGAGCAGAGAGCAGGGCTTAGAAGTAAAGGTTAAAGGA  
GTCATTAAAGCTGGAAAGGAGAAAACTGAGGGATAATTGTGAGCTGTGACTTTTCTCAA  
TATACAAAAGGTTATTTTTAAAACAGGCAACTGAAGAAGAAATGAACAGGCTTGGCTTAC  
[G, A]  
AAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAGGGGAGGCTTGACGGGATCCCA  
ACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTAGATCAATGTGGCTCCAGGGTC  
CAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTCTTGAGAATGGAATGACCTTTGTA  
CTGGTAACATCATTCTTCTGAAACACCTCTCTCCTAGGCCAAAATCCCATGTGCTGAG  
TCCTCGCTCCTGAGCCGGCACTAACGCCCCCTCTCTACCCCCACCTAGGGACATGCCA

FIGURE 3, page 22 of 33

23180 TAAAGGATTCATTAAAGCTTGAAGCTGAAAGGAGAAACTGAGGGATAAATTGTGAGCTGTGACTTT  
TCTCAAATATACAAAAGGTTATTTTTTAAAAACAGGCAACTGAAGAAGAAATGAACAGGCTT  
GGCTTACGAAGAAAGAGCTTGAGGAAGTATAAGGGAAAGTCCCTGAGGGGAGGCTTGACG  
GGATCCCAACCCGAGTGGCCGATGAGACTATTGGGTGGCAGGGGCTAGATCAATGTGGCT  
CCAGGGTCCAGGGCAGCCATGTGATTGTTACTAAGCTGAGATTTCTTGAGAATGGAATGA  
[C, G]  
CTTTGTACTGGTAACATCATTTCTTCTTGAAACACCTCTCTTCTAGGCCAAAATCCCATG  
TCGTGAGTCTCTGCTCCTGAGCCGGCACTAACGCCCCCTCTCTACCCCCACCTAGGGA  
CATGCCAGCCCAGGCTGGGACCTGGTAGAGCGAATGAAAAACAGCCCCGTGAGTACAGG  
CCCCCAGGCCACCCCTGAAAGGTGCCCATCTCCTGCTGGCTGGGGAGGGGACAGCCCCAT  
AAGGGTCCCTCTACCACAGCACTTCTCTGCTTTGGGCTAGCCAAAAGATCCTCGGAGAA

26490 CTCCTCTGCTATAAAGCAAAGCCCTGAGATTGAGCCTGCAAGGACTTACTGAGCACCTA  
CTATGTACCTTGTGTCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCCTA  
CTGGGGAGATGGTTAGAGGAAGAGAACCCTTACACTGAGTCACAGGGGATAGAAGTTAGG  
GGAACACAGGAGAGCAAAACATTTAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGG  
AAAAGGAAGTGTGGCCACCCAGGGCATGGCAAGGGGCTGGAGAAGGCTGAGGTCAGATG  
[A, G]  
CGGATGGGACTGCCAAGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGAGTCC  
ACTGTTGGGTCTGAGATTATGTAGAGCAGGCTGGGGGTTGGGATTGTTTATGGTGTCTAG  
TAGGGGACAAGGGATGATTCTTACAGAGACTCAGCAGCAACAAGAACTGGGCTTCTCAG  
TTTGACCAGGACCACCGAAGCCCTCTGTACCACTCAGTCATTTAGCCAGGCCCCAGA  
GCCCTCCTATGCTCTTGCCATTCTCTCAGAGCGGGCACCAGGGGCTAAAGAGAGTACCTT

26505 AGCAAAGCCCTGAGATTGAGCCTGCAAGGACTTACTGAGCACCTACTATGTACCTTGT  
GCATCACCCAGGATGCTGTGGACACACCTCTAAATCAGCCTCCTACTGGGAGATGGTTC  
AGAGGAAGAGAACCTTACACTGAGTCACAGGGGATAGAAGTTAGGGGAACACAGGAGAGC  
AAACATTTAGGCAGTGGGACCAGCATGGACCAAAGCCCAAAGGAAAAAGGAAGTGTGG  
CCACCAGGGCATGGCAAGGGGCTGGAGAAGGCTGAGGTCAGATGACGGATGGGACTGCC  
[A, G]  
AGAGCCAAGGCCAAAAAGTGGCAGGACCCAGCACTGGCAGAGTCCACTGTTGGGTCTGAG  
ATTATGTAGAGCAGGGTGGGGGTTGGGATTGTTTATGGTGTCTAGTAGGGGACAAGGGAT  
GATTCCTTACAGAGACTCAGCAGCAACAAGAACTGGGCTTCTCAGTTTGACCAGGACCAC  
CGAAGCCCTCTGTACCACTCAGTCATTTAGCCAGGCCCCAGAGCCCTCCTATGCTCT  
TGCCATTCTCTCAGAGCGGGCACCAGGGGCTAAAGAGAGTACCTTTTTTCTTACAGGA

29336 AAATACTTACCCTGCAAATTGAACACCAAGGCCAGGGAAGGGAGTGAGAGACCCCAAAG  
TGGAAGCTGAGAAAATCCCCTTCTCCAGCGGCTAGGCAGCAAGAGATTCCAGAGTAGA  
CTCCTTGTGGTAGGGCCATTCCCCACCCAGAGCCATGTGTAATAATTACTACTCACTTC  
CTCCCCCTCCCTTCATTAACAAAGGCTTAGGCCCCACACAATGGCTCACGTCTGGTG  
TCCCAGTACTCAGGAGGCTGAGATGGGAGGACAGCTTGAGCCAGGAGTTGGAGGCTGC  
[T, C]  
GTTAGCTATGATGATGCCATTGTACTCTGCCTAGACAACAGCGTGAGACCCTATCTCAA  
AAAAAAGAGAAAAAAGAAAAAGGCTTAGCCCTGCCCTACTTAACTCTACCTCAAAT  
TCTCCTTGCCCTCTCTCTGCCCCCTTCCATCTCCCCACCTCCACTCCTGCTTATGTCTCT  
GCCTCTATTGTTCCCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAAGTACCCTCCTT  
CATTCACAAGGCAAGTCTGCTTCCCTCCTTAAGGAGCTTCCCTGCCTGAACCTCACCC

29829 CCCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAAGTACCCTCCTTCATTACAAAGGC  
AAGTCTGCTTCCCTCCTTAAGGAGCTTCCCCTGCCTGAACCTCACCCGCGGACATCTCC  
CCATATCACATTTCAGTCTGTACTTGTAGGGCCCTAAAGCCCCAAAGGGTTCTCATGTTT  
TCACATCTTGGCTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAG  
TCTGTTCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCC  
[C, T]  
GCAAAGTTCATGGGCATTTTGGGAGCTGGTGTGAGATGCTCCCCATCTGACCTGCAGCC  
CCATGTTCTAATTGACCTCTTCTGTCAGTGAGAGAGGGGAGGACTTTGGCCTATGCAAT  
CTGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAAATGGGACTGGGTGGAG  
CTGTGTAGCTAGGGAGCTTCTCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCT  
GAGAACCAGCATAGGGCTTTGAAATGTCCGTGCCATGAATGGGTGGAGAATAAAGATAT

29830 CCTCTCAGGCTCAGGTAGCATTTCCATTCTGCAAAGTACCCTCCTTCATTACAAAGGCA  
AGTCTGCTTCCCTCCTCTAAGGAGCTTCCCCTGCCTGAACCTCACCCGCGGACATCTCCC  
CATATCACATTCAGTCTGTACTTGTAGGGCCCTAAAGCCCCAAAGGGTTCTCATGTTTT



CACATCTTGGCTCATTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAGT  
CTGTTCTCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCCC  
[G,T]  
CAAAGTTCATGGGCATTTTGGGAGCTGGTGTGAGATGCTCCCATCTGACCTGCAGCCC  
CATGTTCTAATTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTATGCAATC  
TGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGGTGGAGC  
TGTGTAGCTAGGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCTG  
AGAACCAGCATAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAAAGTATG

29840 TCAGGTAGCATTTCCATTCTGCAAACCTGACCCTCCTTCATTACAAGGCAAGTCTGCTTC  
CCTCCTCTAAGGAGCTTCCCCTGCCTGAACCTCACCCGCGGACATCTCCCCATATCACAT  
TCAGTCTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCATGTTTTTACATCTTGG  
CTCATTTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACATCTAGTCTGTTCTCTT  
TACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCTCCTCCCGCAAAGTTCA  
[C,T]  
GGGCATTTTGGGAGCTGGTGTGAGATGCTCCCATCTGACCTGCAGCCCCATGTTCTAA  
TTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTATGCAATCTGGTCAGTGG  
CTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGGTGGAGCTGTGTAGCTA  
GGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTGATCCTGAGAACCAGCA  
TAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAAAGTATGTTTGCATCCC

29944 TCTCCCCATATCACATTCACTGTACTTGATGGGCCCTAAAAGCCCCAAAGGGTTCTCA  
TGTTTTTACATCTTGGCTCATTTTTTCCAGATGGATGATAAACTCCTTGAAGATAAGTACA  
TCTAGTCTGTTCTCTTTTACATTCCATGCTTGGGTACTTAAATCCAGCCACCGTGGACTCT  
CCTCCCGCAAAGTTTATGGGCATTTTGGGAGCTGGTGTGAGATGCTCCCATCTGACCT  
GCAGCCCCATGTTCTAATTGACCTCTTCGTGCAGTGAGAGGAGGGGAGGACTTTGGCCTA  
[C,T]  
GCAATCTGGTCAGTGGCTCAGACCCAGCCTTTCAGGCAGAGGCTTTGGAATGGGACTGGG  
TGGAGCTGTGTAGCTAGGAGCTTCTCCCACCAGGAGCCGCTGGGTTCAACTCATCTCTG  
ATCCTGAGAACCCAGCATAGGGCTTTGAAATGTCCGTGCCCATGAATGGGTGGAGAATAAA  
AGTATGTTTGCATCCCACTAGAGTAGCCCTTAAAGTCACTGTCTTTAGGGTGAGTTGA  
CTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGTCTGTGCAGCCTTGCCAG

30468 CCTTTAGGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGT  
CTGTGCAGCCTTGCCAGGAGGGTTGAGCAGCTTCTCTCTGTCCCAGCATGGCATCTC  
CAGTTTCTCTACTGGCACCATAACACACAGCGTGAGGACTTTGCGGTTGTGGTGCAGCC  
TTTCTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTTAGGGAGGC  
TCACGTATGGGGGCCCTTATCACAGACGATGGATGTATTTCTTCTCTAAGTGGGCTTTTT  
[T,-]  
TTTTTTTTTAACCATCTCTCTCCAAGAGGATTCTGAGGGTGGCTTTTTCCACATTACCTC  
CTTTTTGTGGGGGCTGGGCTGTGATTGGAACCTCAGATGTACTTTGAAAGGAAATCAATAG  
TGACTAAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAGCTTCT  
CTAAAAATGTTACTCTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACT  
ATGGGTGGTCCAATTCTTGTCTGTTTAAAAAGAAAATCCGGCCGGGTGCAGTGGCTCA

30471 TTAGGGTGAGTTGACTCCCGTCAACAACCAATCCAAGGCAGCAGGACTGGACCCTGTCTG  
TGCAGCCTTGCCAGGAGGGTTGAGCAGCTTCTCTCTGTCCCAGCATGGCATCTCCAG  
TTTCTCCTACTGGCACCATAACACACAGCGTGAGGACTTTGCGGTTGTGGTGCAGCCTT  
CTTCCAAAACACACTCACCCCACTGAACGAGGTGAGCTGCAGGTATTTTAGGGAGGCTCA  
CGTATGGGGGCCCTTATCACAGACGATGGATGTATTTCTTCTCTAAGTGGGCTTTTTTTT  
[T,-]  
TTTTTTAACCATCTCTCTCCAAGAGGATTCTGAGGGTGGCTTTTTCCACATTACCTCCTT  
TTTGTGGGGGCTGGGCTGTGATTGGAACCTCAGATGTACTTTGAAAGGAAATCAATAGTGA  
CTAAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAGCTTCTTA  
AAAATGTTACTCTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATG  
GGTGGGTCCAATTCTTGTCTGTTTAAAAAGAAAATCCGGCCGGGTGCAGTGGCTCATGC

30802 TGAGGGTGGCTTTTTTCCACATTACCTCCTTTTTTGTGGGGGCTGGGCTGTGATTGGAACCTC  
AGATGTACTTTGAAAGGAAATCAATAGTACTAAGCTCCCAGGCCTGGCCCTGATGTTTT  
CTGGATTGGGATAGAATGGAAGCTTCTTAAAAATGTTACTCTTTTCAACTCTTAGGATA  
GGGGTGCTGAAAGAAAAGGGAGAGACTATGGGTGGGTCCAATTCTTGTCTGTTTAAAAAG  
AAAATCCGGCCGGGTGCAGTGGCTCATGCCTGTAATCTCAGCCTTTGGGAAGCCAAGGC  
[G,A]

FIGURE 3, page 24 of 33

GTTGAATCACGAGGTTAGGAGTTTGAAGACCAGCCTGGCCAACATGGTGAAACCCCGTTTC  
TACTAAAAATACAAAAAGTTAGCTGGGCGTGGTGGCAGGCACCTGTAATCCCAGGTACTC  
GGGAGGCAGAAGTTGCAGTGAGCTGAGATTATGCCACTGCACTCCAGCCTGGCTGACAGT  
GCGAAACTCCGTCTCAAAAAAAGAAAAAAGAAAAAAGAAATTTCTAAATTTCTGGGAGT  
TTTTCCATCAGTATCTGAGCAAGTTGGCAGGAAAGTTGAAAGAATGAAAGGAGACATGCC

30894 AAGCTCCCAGGCCTGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTCCTAAA  
AATGTTACTCTTTTTCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATGGG  
TGGGTCCAATTCTGTCTGTTTAAAAAGAAAATTCGGCCGGGTGCAGTGGCTCATGCCT  
GTAATCTCAGCCTTTGGGAAGCCAAGGCGGTTGAATCACGAGGTTAGGAGTTTGAGACCA  
GCCTGGCCAACATGGTGAAACCCCGTTTCTACTAAAAATACAAAAAGTTAGCTGGGCGTG  
[G, A]  
TGGCAGGCACCTGTAATCCCAGGTACTCGGGAGGCAGAAGTTGCAGTGAGCTGAGATTAT  
GCCACTGCACTCCAGCCTGGCTGACAGTGCGAAACTCCGTCTCAAAAAAAGAAAAAAG  
AAAAAAGAAATTTCTAAATTTCTGGGAGTTTTTCCATCAGTATCTGAGCAAGTTGGCAGGA  
AAGTTGAAAGAATGAAAGGAGACATGCCAGGGCACCTGCTGGGAGAGTGAGTGGGGCTC  
AGGTAGCAGAGCCCTTTCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGAC

30907 TGGCCCTGATGTTTTCTGGATTGGGATAGAATGGAAAGCTTCCTAAAAATGTTACTCTTT  
TCAACTCTTAGGATAGGGGTGCTGAAAGAAAAGGGAGAGACTATGGGTGGGTCCAATTCT  
TGTCTGTTTAAAAAGAAAATTCGGCCGGGTGCAGTGGCTCATGCCTGTAATCTCAGCCT  
TTGGGAAGCCAAGGCGGTTGAATCACGAGGTTAGGAGTTTGAGACCAGCCTGGCCAACAT  
GGTGAAACCCCGTTTCTACTAAAAATACAAAAAGTTAGCTGGGCGTGGTGGCAGGCACCT  
[G, A]  
TAATCCCAGGTACTCGGGAGGCAGAAGTTGCAGTGAGCTGAGATTATGCCACTGCACTCC  
AGCCTGGCTGACAGTGCGAAACTCCGTCTCAAAAAAAGAAAAAAGAAAAAAGAAATTT  
CTAAATTTCTGGGAGTTTTTCCATCAGTATCTGAGCAAGTTGGCAGGAAAGTTGAAAGAAT  
GAAAGGAGACATGCCAGGGCACCTGCTGGGAGAGTGAGTGGGGCTCAGGTAGCAGAGCC  
CTTTCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGACACTGACCTCACCT

31447 CCTTTCCAGGATGATAACCTCCTTGCCGTTGGTTGCAGAGAGGGGACACTGACCTCACCC  
TTCTTCTCCGAGGACTGTTTTCACTTCTCAGACCGCGGGCATGCCGAGATGGCCATCGCA  
CTCTGGAACAACATGGTGAGCAGCCAAGGGCCTGGTGGGCCTTGTCAAGGGGGATCTAA  
GGATATTGACACTCTGTCTCACAATGGCAAACTACTGGAGACATGGCTCCTTTCTCCCC  
AAAGCCCAAAGTGGCAGCACACCTTATTGGTCCTGATAGATTAATTCCAAAGGGAAAATA  
[C, A]  
CCTATATTTATCCAACACCCCTTTGAAAGTTATACAAACACACACTCACACAACCTTTATTC  
TTTGTTCCTTCAGCAATGCCAGGTACTGCGAGGGGATCCCTTTGTAATCAGATAGGTTG  
GCTAGATGAAAATACCAACTTCTACCTCGTACTGTGTGACCTTGGGCAAACGATCTCTCT  
GGCCACCTGTATCAACATCTATAAAACAGTGAAAACAGACAGGTCTCAGACAACGCATT  
GAGATCATGTGTACATGGCACCTAGCACAAATAGTTAGCACTCAGCAAATGTCACCAACAT

31603 GGGCCTTGTCAGGGGGGATCTAAGGATATTGACACTCTGTCTCACAATGGCAAACTAC  
TGGAGACATGGCTCCTTTCTCCCCAAAGCCCAAGTGGCAGCACACCTTATTGGTCCTGA  
TAGATTAATTTCCAAAGGGGAAATACCCCTATATTTATCCAACACCCCTTTGAAAGTTATACA  
AACACACACTCACACAACCTTTATTCTTTGTTCTTTCAGCAATGCCAGGTACTGCGAGGG  
GATCCCTTTGTAATCAGATAGGTTGGCTAGATGAAAATACCAACTTCTACCTCGTACTGT  
[G, A]  
TGACCTTGGGCAAACGATCTCTTGCCACCTGTATCAACATCTATAAAACAGTGAAAAC  
AAGACAGGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAGCACAAATAGTTA  
GCACTCAGCAAATGTCACCAACATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCCA  
ACTCATTTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATA  
CCTATTCCAGTCTTTCTGATGGGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGG

31685 CCAAAGCCCAAAGTGGCAGCACACCTTATTGGTCCTGATAGATTAATTCCAAAGGGAAAA  
TACCCTATATTTATCCAACACCCCTTTGAAAGTTATACAAACACACACTCACACAACCTTA  
TTCTTTGTTTCTTCAGCAATGCCAGGTACTGCGAGGGGATCCCTTTGTAATCAGATAGG  
TTGGCTAGATGAAAATACCAACTTCTACCTCGTACTGTGTGACCTTGGGCAAACGATCTC  
TCTGGCCACCTGTATCAACATCTATAAAACAGTGAAAACAGACAGGTCTCAGACAACGC  
[A, G]  
TTGAGATCATGTGTACATGGCACCTAGCACAAATAGTTAGCACTCAGCAAATGTCACCAAC  
ATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCAACTCATTTCTCTAAACATCGAA  
AAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATACCTATTCCAGTCTTTCTGATG

FIGURE 3, page 25 of 33

GGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGGGGTCTTTCAGGGACACCTGCCT  
GGTGCTTCCACTCTGCCTTCTGTGGCTGGCCACCAGCAACTGAACGGTTTCCGCACAGCA

31833 ACTGCGAGGGGATCCCTTTGTAATCAGATAGGTTGGCTAGATGAAAATACCAACTTCTAC  
CTCGTACTGTGTGACCTTGGGCAAACGATCTCTCTGGCCACCTGTATCAACATCTATAAA  
ACAGTGAAAACAAGACAGGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAG  
CACAATAGTTAGCACTCAGCAAATGTCACCACCATCAGCCTTCCAAGCACTCCGGGGCTCA  
ACTCATACCCAACCTCATTTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTT  
[C, T]  
GAGGCTGATACCTATTCCAGTCTTCTGTATGGGAAGAAGGGACCTTATGAAATGAACAT  
ACAGTCTGGGGGTCTTTCAGGGACACCTGCCTGGTGCTTCCACTCTGCCTTCTGTGGCTG  
GCCACCAGCAACTGAACGGTTTCCGCACAGCACTTGACCTGTACCCCCAACAACTGGAT  
CCTCTTGACGGAGCAAATGAAATGCCTTCCCAACCAATGGTTTCTTTTAAATCCAGGCT  
CAGTGGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCCTTTGTCCTATGACAACT

31970 GGTCTCAGACAACGCATTGAGATCATGTGTACATGGCACCTAGCACAATAGTTAGCACTC  
AGCAAATGTCAACACCATCAGCCTTCCAAGCACTCCGGGCTCAACTCATACCCAACCTCAT  
TTCTCTAAACATCGAAAAGTGGAGATCCACACAGCCTGTTTTCCGAGGCTGATACCTATT  
CCAGTCTCTTCTGTATGGGAAGAAGGGACCTTATGAAATGAACATACAGTCTGGGGGTCTT  
TCAGGGACACCTGCCTGGTGCTTCCACTCTGCCTTCTGTGGCTGGCCACCAGCAACTGAA  
[C, T]  
GGTTTCCGCACAGCACTTGACCTGTACCCCCAACAACTGGATCCTCTTGACGGAGCAA  
ATGAAATGCCTTCCCAACCAATGGTTTCTTTTAAATCCAGGCTCAGTGGGTAACACAATC  
CCCACCCCAACCTGTATGTTCCCTCCTTTGTCTATGACAACTAAACAAGCTACATTCCA  
GCTCCTTTTATCACAGTTTCAGGCCCGTAGTGTCTCTGCCAACCCCGCTGTGCAAACGT  
TCCCACCCCTGTACAGTCTATCCAGTATGTCCAGCATCCCACTCGGCTGACTCACAATATT

32177 CCTTATGAAATGAACATACAGTCTGGGGGTCTTTCAGGGACACCTGCCTGGTGCTTCCAC  
TCTGCCTTCTGTGGCTGGCCACCAGCAACTGAACGGTTTCCGCACAGCACTTGACCTGTG  
ACCCCAACCACTGGATCCTCTTGGCAGGAGCAAATGAAATGCCTTCCCAACCAATGGT  
TTCTTTTAAATCCAGGCTCAGTGGGTAACACAATCCCCACCCCAACCTGTATGTTCCCTCC  
TTTGTCTATGACAACTAAACAAGCTACATTCCAGTCTCTTTTATCACAGTTTCAGGCC  
[A, G]  
TAGTGTCTCTGCCAACCCCGCTGTGCAAACGTTCCCACCCCTGTACAGTCTATCCAGTAT  
GTCCAGCATCCCACTCGGCTGACTCACAATATTGACTTTCTCCTTAGCTATACCATCTCC  
TCCTCTCTAGCAACCTCTTCTTTAAGAACAGCATGTAACTGGCTTTATCCTTGGCCTA  
GTTAATGGCAGACTCAGCTTATGTGCACTTCCATTGTACAGGGGGTTTTCCTCCTGTGGAC  
ATCACGTACCTGCCCACTCCAAGAAGTCTATTGTACTCTTTTACGCCCAAGACTCCGGAT

33018 AAGGTATGGCCTTCCCTACCAGGTGGCACTCCAAGTCTGCTTAAATCTGGGACCTCCAGG  
AATCTCCTGGGGCTGGATAGCCATAGTGACGGCTGGAACATGAAAAAGAGTCCATTGGTT  
TCTTTTCTTGTGAATTAACAATGTAGCTCTGGCCAGGCACGGTGGCTCATGCCTGTAATC  
CCAGCACTTTGGGAGGCCGAGGCAGGTGGATCGCTTGAGCCAGGAATTAGACACCAACC  
TGGGCAACACAGGGGAGATTCTGTCTCTACAAAATAATCAAAATATTAGCCAGGTGTGG  
[T, C]  
GGTGCATGCCTGTAGTCCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTTGAGCATG  
GGAGGTCAAGGCTGCAATGAGCCGAGATGGCACCACCGCACTCCAGCCTGGGCAATAGAG  
TGAGACCCTATATCTCAAAAAACAAATAGAAAAAATAATATGTAGCTCTGGCCTTCT  
CTTCTAAAGCAGTTTCAGTAGCTCTTCCATTACCCAGGTAAGAGGCCTTTATTTTCATAA  
AGATAAGTGGGAGGAGTTTAGATATGAAAACAAACGTAAACACCGCACTGGAGCTATTG

33090 CTGGATAGCCATAGTGACGGCTGGAACATGAAAAAGAGTCCATTGGTTTCTTTTCTGTG  
AATTAACAATGTAGCTCTGGCCAGGCACGGTGGCTCATGCCTGTAATCCAGCACTTTGG  
GAGGCCGAGGCAGGTGGATCGCTTGAGCCAGGAATTAGACACCAACCTGGGCAACACAG  
GGGAGATTCTGTCTCTACAAAATAATCAAAATATTAGCCAGGTGTGGTGGTGCATGCCT  
GTAGTCCCAGCTGCTCAGAAGGCTGACGTGAGAAGATCACTTGAGCATGGGAGGTCAAGG  
[T, C]  
TGCAATGAGCCGAGATGGCACCACCGCACTCCAGCCTGGGCAATAGAGTGAGACCCTATA  
TCTCAAAAAACAAATAGAAAAAATAATATGTAGCTCTGGCCTTCTCTTCTAAAGCAG  
TTCAGTAGCTCTTCCATTACCCAGGTAAGAGGCCTTTATTTTCATAAAGATAAGTGGGA  
GGAGTTTAGATATGAAAACAAACGTAAACACCGCACTGGAGCTATTGTGGAACAAAAC  
AAGACTGTCCATGGTTCCCAGCCATTATTATCTCAGCCATACCCGAATTTCAAAATAA

FIGURE 3, page 26 of 33

33993 ATTACAGACCAGGTTTCTAGTCCTTCTCCTGGTGACCTGGGCATGCCACCACCCTCCCC  
ACTGCTCCCAACCTGATAAGCACATATATACCGGTGAATTCATGTCTCACAATTAGAGT  
CCTATGACATAGTGTCTGCAGGCTTTGGCTGATGTTCCCATAGTGTCTGCAGGCTTTGGC  
TGATGTTCCAGGGTTCCCTACTAGGAAGCAAAAGCACCTTAACTATTTTCATCTTATT  
TCATCTCCTGCCCTCCTCTCACGTCTTCTCGAGACTTTTGCAAAGGCAAAGCCAGAAG  
[C, T]  
TCCAGCAGCACCAGGGGATATTTTCTCTTCTCTGCCTTCTTCTGTCTTCTTATCTGAA  
GAAGTTTCTCTTTCCCGAGGCCTAGTCCTCTACTGCTGCCTCTACTCCCTCTTCTGCAGA  
AATCCTGCTCTCAGCCAGTGTGTTGTATCTCCCAGGTGCTGGGTGACAGCTCCAGCCTCC  
TAACCTGACATCCCTGTCTTACAGACTTAGAGCTCTTAGAATCGTGACTCTCAGCTCTGGCT  
GCATATTAGAATCATTCAGGGACATTGTGTATGTGTGTATGTATGTATATATGTATGA

34284 AGCCAGAAGCTCCAGCAGCACCAGGGGATATTTTCTCTTCTCTGCCTTCTTCTGTCTT  
CTTATCTGAAGAAGTTTCTCTTTCCCGAGGCCTAGTCCTCTACTGCTGCCTCTACTCCCT  
CTTCTGCAGAAATCCTGCTCTCAGCCAGTGTGTTGTATCTCCCAGGTGCTGGGTGACAGC  
TCCAGCCTCCTAAGTACATCCCTGTCTTACAGACTTAGAGCTCTTAGAATCGTGACTCTC  
AGCTCTGGCTGCATATTAGAATCATTCAGGGACATTGTGTATGTGTGTATGTATGTATGT  
[G, A]  
TATGTATGAATGTGTGTGTATGTGTGTGTGTGTATGTATGTATGTATGTGTATGTATGTA  
TGTATGTATGTATGACAGAGTCTCACTCTGTTGCCAGGTTGGAGAGCAATGGCACCATC  
TCAGTTCACTGCAACCTCCGTCTCCTGGATTCAAGCGATTCTCCTGCCTCAGTCTCCCAA  
GTAGCTGGGGTTATAGGTGCATGCCACCATGACCAGCTAATTTTGTATTTTAGTAGAG  
ACAAGGTTTCGCCATGTTGGCCAGGCCAGGCTGGTCTTTAACTCCTGACCTCAGGAGATC

34314 TTTTCTCTTCTCTGCCTTCTTCTGTCTTCTTATCTGAAGAAGTTTCTCTTTCCCGAGG  
CCTAGTCTCTACTGCTGCCTCTACTCCCTCTTCTGCAGAAATCCTGCTCTCAGCCAGTG  
TTTGTATCTCCCAGGTGCTGGGTGACAGCTCCAGCCTCCTAAGTACATCCCTGTCTTC  
AGACTTAGAGCTCTTAGAATCGTGACTCTCAGCTCTGGCTGCATATTAGAATCATTCAGG  
GACATTGTGTATGTGTGTATGTATGTATGTATATATGTATGAATGTGTGTATGTGTGT  
[G, A]  
TGTATGTATGTATGTGTATGTGTGTATGTATGTATGTATGTATGTATGACAGAGTCTCACTCTG  
TTGCCAGGTTGGAGAGCAATGGCACCATCTCAGTTCACTGCAACCTCCGTCTCCTGGAT  
TCAAGCGATTCTCCTGCCTCAGTCTCCAAGTAGCTGGGGTTATAGGTGCATGCCACCAT  
GACCAGCTAATTTTGTATTTTAGTAGAGACAAGGTTTCGCCATGTTGGCCAGGCCAGG  
CTGGTCTTTAACTCCTGACCTCAGGAGATCCACCACCTCGGCCTCCCAAAGTGCTGGGA

35392 GACGTGAGCATGATGGTGTAGAGACTCACTGGGGGATGAATAGTCTGGAAGAAGGTGGA  
AAGGGGCTTTTGAGGACTATAATAGTCTGTTGCCTGACTGGATGCTGGTATGTTTATTTT  
ATCGAACTTATCTGTTGCTCACTTATGATTTGTACTCGTTTCTATGTGTATGTAGCTT  
CAATTAAGTTTTACTTGAGGCCGGGTACAGTGGCTCACACCTGTAATCCCAGCACTTTG  
GGAGGCCGAGGCAGGCAGATCCCTGAGGTGAGGAGTTCAATACCAGCCTAGCCAACATG  
[A, G]  
TGAAACCCCATCTCTACTAAAAATACAAAATTAGCCAAGCGTGGTGGCACGTGCCTATAA  
TTCCAGCTACTTGGGAGGCTGAGACAGGAAAATCGCTTGAAACCAGGAGGCAGGGGTTGC  
AGTGAGCCAAGATTGCATCATTCAGCTCCAGCCTGGGTGACAAGAGTAAACTCTGTCTC  
AAATTTAAAAAAGTTTACTTGAAAAACAATATCAGTGCCTGACCGG  
GCTTATCCCGAGAGTCTGACTTAATTGGTCTGGAGTGCAGCTGGATTGGTACTTTG

35599 ACAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGGCCGAGGCAGGCAGATCCCTGA  
GGTCAGGAGTTCAATACCAGCCTAGCCAACATGATGAAACCCATCTCTACTAAAAATAC  
AAAATTAGCCAAGCGTGGTGGCACGTGCCTATAATTCCAGCTACTTGGGAGGCTGAGACA  
GGAAAATCGCTTGAAACCAGGAGGCAGGGGTGCACTGAGCCAAGATTGCATCATTCAC  
TCCAGCCTGGGTGACAAGAGTAAACTCTGTCTCAAATTTAAAAAAGGCTGACCGG  
[T, A, G]  
TTTACTTGAAAAACAATATCAGTGCCTGACCGGGCTTATCCCGAGAGTCTGACTTAAT  
TGGTCTGGAGTGCAGCTGGATTTCGGTACTTTGTGAAAGCTCCTGAGATTATTTTATGT  
GCAGGGTTTATGAACCGCTGCCTTAGATCTGGTCCCCACAGAGAAATCAAGTAATCTGTA  
TAAAGAAAACCTGACCCAGTCACTCCCTGCTTTCAAACCTCCAAAGCCTCCACCTCT  
GAAGGAGGCAGGCCAGGCCCATAGCACAGCACACTAGGCCTCTGGGACTTGGCCTGGTT

35997 AGCTCCTGAGATTATTTTAAATGTGCAGGGTTTATGAACCGCTGCCTTAGATCTGGTCCCC  
ACAGAGAAATCAAGTAATCTGTATATAAAGAAAACCTGACCCAGTCACTCCCTGCTTTCA  
AACTTCCAAAGCCTCCACCTCTGAAGGAGGCAGGCCAGGCCCATAGCACAGCACACTA

FIGURE 3, page 27 of 33

GGCCTCTGGGACTTGGCCTGGTTACCTGATTAACTCTCTGGCTACCATTTCCACCAGC  
GTCTGCCTCGCATGTTACAGTCTAGTGACTCCAGCAGCGTCCTGCACCACCTGTGGTGTT  
[C, A]  
CACACCTCTGCTAACTCTTGCTCTCCTCCTTCTCCTGGATTGCCCTTCTCACCTCCTTGC  
CCACTCCACCACCTCAACTCAGGTGCCACCTCCTGCAGGAAGCTACCTCTGAATCTCCAGG  
ACAGGCCAGTGGCCACCCAGGTCCATTACACCTGCCCAGTCCTGTCAATTGCTACGTG  
GTTGGTAGCCACAGTGCCTGGCTTAGGAAAGACTGGTTCTAGGAAAAACAATTTCAATCC  
CTGTGGCCAGCTCCAAGCCTTCCCCCGCCAAGCTTCTCCATTCAGGTCTCTGTGAATTTA

36085 GAAAACCTGACCCAGTCACTCCCCTGCTTTCAAACCTTCAAAGCCTCCCACCTCTGAAGG  
AGGCAGGCCAGGCCCATAGCACAGCACACTAGGCCTCTGGGACTTGGCCTGGTTACCT  
GATTAACTCTCTGGCTACCATTTCCACCAGCGTCTGCCTCGCATGTTACAGTCTAGTGA  
CTCCAGCAGCGTCCTGCACCACCTGTGGTGTTCACACCTCTGCTAACTCTGCTCTCCT  
CCTTCTCCTGGATTGCCCTTCTCACCTCCTTGCCCACTCCACCACTCAACTCAGGTGCCA  
[C, T]  
CTCCTGCAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCACCCAGGTCCATT  
ACACCTGCCCAGTCTGTCAATTTGCTACGTGGTTGGTAGCCACAGTGCCTGGCTTAGGA  
AAGACTGGTTCTAGGAAAAACAATTTCAATCCCTGTGGCCAGCTCCAAGCCTTCCCCGC  
CAAGCTTCTCCATTCAGGTCTCTGTGAATTTAATTAATTCATCCATCCATCAAACAAGTA  
TTTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGCAGTGAAA

36270 GCAGCGTCTGCACCACCTGTGGTGTTCACACCTCTGCTAACTCTGCTCTCCTCCTTC  
TCCTGGATTGCCCTTCTCACCTCCTTGCCCACTCCACCACTCAACTCAGGTGCCACCTCC  
TGCAGGAAGCTACCTCTGAATCTCCAGGACAGGCCAGTGGCCACCCAGGTCCATTACAC  
CCTGCCCAGTCTGTCAATTTGCTACGTGGTTGGTAGCCACAGTGCCTGGCTTAGGAAAGA  
CTGGTTCTAGGAAAAACAATTTCAATCCCTGTGGCCAGCTCCAAGCCTTCCCCCGCCAAG  
[C, T]  
TTCTCCATTCAGGTCTCTGTGAATTTAATTAATTCATCCATCCATCAAACAAGTATTTAC  
TGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGCAGTGAAAAGATG  
ACTGTACTCTCATGGGACATACAGGATAGTAGGAAAAAGACAGATAATCAACAAGGTCA  
TTTCTGACCACATCTGTGGTTTAAAGAAAAAGTCAAGCAGAGTGATGTGATACAGAGTAAT  
GGTGGGGGAGAGGGAGGCCTCCCTGAAGAAGTGACAGTGAATTGAGAAGCGCATGTCAAG

36481 GGTAGCCACAGTGCCTGGCTTAGGAAAGACTGGTTCTAGGAAAAACAATTTCAATCCCTG  
TGGCCAGCTCCAAGCCTTCCCCCGCCAAGCTTCTCCATTCAGGTCTCTGTGAATTTAAT  
AATTCATCCATCCATCAAACAAGTATTTACTGAGCACTAATATGTGCTAGGTACTGCTCC  
AGGTGCTGAGGACTCAGCAGTGAAAAGATGACTGCTACTCTCATGGGACATACAGGATAG  
TAGGGAAAAGACAGATAATCAACAAGGTCAATTTCTGACCACATCTGTGGTTTAAAGAAAA  
[G, A]  
TCAAGCAGAGTGATGTGATACAGAGTAATGGTGGGGGAGAGGGAGGCCTCCCTGAAGAAG  
TGACAGTGAATTGAGAAGCGCATGTCAAGGGGTTGCCAGGCAGAGGAAATAGGACCCACA  
TGGGCCCTAGAGTCAGGAGTGAGCTTGAAGTGTCTGAGGAACTTAAAGGCCAATGTGACCA  
GAGGGAAGTGAAACAAGGTGAAAAAGTTGGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGC  
AGTAGAGTGATATGCTCTGGCTTACCCCTGGGTCCGTGTACCCTGGACTGGAAGAAAGCA

36619 ACAAGTATTTACTGAGCACTAATATGTGCTAGGTACTGCTCCAGGTGCTGAGGACTCAGC  
AGTGAAAAGATGACTGCTACTCTCATGGGACATACAGGATAGTAGGGAAAAGACAGATAA  
TCAACAAGGTCAATTTCTGACCACATCTGTGGTTTAAAGAAAAAGTCAAGCAGAGTGATGTG  
ATACAGAGTAATGGTGGGGGAGAGGGAGGCCTCCCTGAAGAAGTGACAGTGAATTGAGAA  
GCGCATGTCAAGGGGTTGCCAGGCAGAGGAAATAGGACCCACATGGGCCTAGAGTCAGGA  
[G, A]  
TGAGCTTGAAGTGTCTGAGGAACTTAAAGGCCAATGTGACCAGAGGGAAGTGAACAAGGT  
GAAAAAGTTGGGCAGGGGCCAGGTCCCTAGATGCTTCTAAGCAGTAGAGTGATATGCTCT  
GGCTTACCCCTGGGTCCGTGTACCCTGGACTGGAAGAAAGCAAGGGTGGACCTGGAAGA  
CCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGCTGAGAGTAGGGTCAGGGC  
AGAGGAGGAGAGACGCTGTCTGGGCTGGCGGATGGATGATGGGGAAGAGGAACAAAGGA

37088 GACCTGGAAAGACCACTAGGAGGCTGCTGTTGATGGGTGAGAGAGGAAGGGGGCTGAGAG  
TAGGGTCAGGGCAGAGGAGGAGACGCTGTCTGGGCTGGCGGATGGATGATGGGGAAG  
AGGAACAAAGGATGACTTTTGGTTTGGGGTCTAAGAACTGGGTGGATGATTGAGCAGG  
TAGAGAAAAAATCAGCGTGGGAGGAAAAAATCAAGACTTCTGTTTTGGACATGGTGCA  
AACTGCCCTTCCAGACATCCACATAGAGGTATCAGGATACAGAAGTTTGGAACTCACAGAG  
[G, C]

FIGURE 3, page 28 of 33

AAGTCAAGGCTGGAGATTGAAAAAAAAAAAAAAAAAAAAAGTGGGGTTATTAGCATA  
GAGGGCCAATATGGTGAAACCCGTGTCTCTACTGAAAATACAAAATTATCCAGGCATGGT  
GGCATGCACCTGTAAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGAACCC  
AGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCATTGCACTCCAGCCTGGGTGA  
CAGGGCAAGATTCCATCTAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGCTCCTA

37204 GAAGAGGAACAAAGGATGACTTTTTGGTTTGGGGTCTAAGAACTGGGTGGATGATTGAG  
CAGGTAGAGAAAAATCAGCGTGGGAGGAAAAAAATCAAGACTTCTGTTTTGGACATGG  
TGCAACTGCCTTCCAGACATCCACATAGAGGTATCAGGATACAGAAGTTTGAAGTCA  
AGAGGAAGTCAAGGCTGGAGATTGAAAAAAAAAAAAAAAAAAAAAGTGGGGTTATTA  
GCATAGAGGGCCAATATGGTGAAACCCGTGTCTCTACTGAAAATACAAAATTATCCAGGC  
[G, A]  
TGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCTGAGGCAGGAGAATTGCTTGA  
ACCCAGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCATTGCACTCCAGCCTGG  
GTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGCT  
CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGCCAACTACCTAGCCCTGGGCAT  
TCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAA

37485 ATACAAAATTATCCAGGCATGGTGGCATGCACCTGTAATCCCAGCTACTCAGGGAGGCT  
GAGGCAGGAGAATTGCTTGAACCCAGAGATGGGGTGGAGGTTGCAGTAAGCTGAGATCGT  
GCCATTGCACTCCAGCCTGGGTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAGCCAC  
TACAGGATCAACTAAGAGCTCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGCC  
AACTACCTAGCCCTGGGCATTTCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGA  
[G, A]  
TGTGGAGGTAGATGGGAAATGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGT  
GTTTCAAGAGAGACAGAGCTGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTG  
AAGACAAAAAAAAAAAAATGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAG  
GGCCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGA  
AGGAGAGATTAGGAAGTGTGCCAGCACCTATAGACAGCTCTTCCAGAAGTTATGAGAA

37624 GGTGACAGGGCAAGATTCCATCTAAAAAAAAAAAAAGCCACTACAGGATCAACTAAGAGC  
TCCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGCCAACTACCTAGCCCTGGGCA  
TTCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAA  
ATGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGTGTTCAGAGAGACAGAG  
CTGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAA  
[-, A]  
TGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAG  
CGATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGC  
TGCCAGCACCTATAGACAGCTCTTCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTG  
GAGGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGGAGATGTCCGAGCAGGTTGTG  
TACTGACGAGAAGGAACCAAGTAGAAAGGGAGAACTGATGCACTCATCAAACCTTGTAA

37685 CCTAGAGAAAGAATAGGTAGGTAGAAAAGAGTGTAAAGCCAACTACCTAGCCCTGGGCAT  
TCATTCCAGCTTTCAACTCCAGTGAGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAA  
TGAGAAACAATGCTGTGTCCAGAGAGCTAAGAGAAGTCAGTGTTCAGAGAGACAGAGC  
TGTCAACTTTGATGGATGCTTCTGAGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAA  
TGATCTTTGGCTCTGCCCATATGGCGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAG  
[C, T]  
GATGGAGACTGCAGACTGGCTGGAGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGT  
GCCAGCACCTATAGACAGCTCTTCCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGG  
AGGGGACATGGATCAAAGAAAGGGCAGGTGAAGGAGGGGAGATGTCCGAGCAGGTTGTGT  
ACTGACGAGAAGGAACCAAGTAGAAAGGGAGAACTGATGCACTCATCAAACCTTGTAA  
CACGATCATCTTCTGTGTGAATTAGTTCTGGGTTCTGGAATAGCATCGGAATCAGCCG

37769 AGAGATGAGAAGGAGAGTGTGGAGGTAGATGGGAAATGAGAAACAATGCTGTGTCCAGAG  
AGCTAAGAGAAGTCAGTGTTCAGAGAGACAGAGCTGTCAACTTTGATGGATGCTTCTG  
AGAAGCCAAGCAAGTTGAAGACAAAAAAAAAAAAATGATCTTTGGCTCTGCCCATATGG  
CGATCGTTGGTGGCCAGGGCCAGAGCTTCCATCCAGCGATGGAGACTGCAGACTGGCTGG  
AGCGAGCAGCAGAGAGAAGGAGAGATTAGGAAGTGTGCCAGCACCTATAGACAGCTCTT  
[C, T]  
CCAGAAGTTATGAGAAGTAACAGCCACGGTCACTGGAGGGGACATGGATCAAAGAAAGGG  
CAGGTGAAGGAGGGGAGATGTCCGAGCAGGTTGTGTACTGACGAGAAGGAACCAAGTAGAA  
AGGGAGAACTGATGCACTCATCAAACCTTGTAAATCAGCATCATCTTCTGTGTGAATTA

FIGURE 3, page 29 of 33

GTTCTGGGTTCTGGAATAGCATCGGGAATCAGCCGCGCTGACCTTTAGCATTTATTCTG  
TCACTGTTACGATAGACTTGAGTTTCCTCAGTTCTTAAGAAAGTGAAATAATACTACCT

38897 AGAGGGTCTGGGATCAAAGGTATTTACCCAGGGATATTTAGATAAATCTTTTCATC  
TATGTGGAACATACAAAGTGGCGCAAGTGAGAACTCCGATTTCTAAGGTTGACAAG  
TCAAGTGCAGTAATGATGTGATGTAACCAATATGTTTCCAACTTTCTAAGGTTGACT  
AGCCCCATGCACCTTTGAGAAGTTGGTAAATAGGATTGTCGTGTTTATAAAATTGAAAA  
CACGGTGTCTTGCAATCACAGCCACTCACAAAGGAAGCCAGAGATGGTCCCAGCCCCCTC  
[G, A]  
CAGACTTCTGTGGACTCAGGACTGGTGGTCTCTCCTGGGCCTTGCTGTACCCGGCAAAT  
CCAGGGGCACAGACTCAGGGTCTGCCCTGCCGACAGATGCTGCCAGCCTTCTGTGTGT  
CATAAGTCAACTCCCCTCAGCCCCAGGCTGCTGGGTCCCTGCTGTGGGCCAAAAACCAG  
CCACTTCGTGTTTCTATCCCCACCCGTTCCCGAGGGAGGGGCTCTGGTGTGAGACA  
CCCCCTCAGAGAGGAAAGTGTCTCCAGCTTTGGAGAGAATCGAGGTGTCTTTCTCTCT

40155 AGACTGAGTGACAGAGTGAGACTGTCTCAAAAATAAAGTGCATCAAGCAGCTGTCCCGTG  
CCAGGCAGTATACTAGGATCTGGGGATCGGGAGGCAAAGATAAAATAGACTCAGTGTCTG  
TTCCTGGAGCCTGCAATGGTCTTCTCCCTCGCCACACCCACTGCCCTTGCTGGCCCCAC  
CTTCGAAGCCTGTGACTGTCTCCCAGCTCTCTCTCCCTTCTCCATCCACCCTACA  
CTTGCTGCCAGACACAGATAGACCTTCTGGAATAAAGTGTGCCCCATCAAGGCTGCTTGA  
[A, G]  
ATCCTTGCTGATCCCTACTGCCCATTGACCAGAGTCTGGAGGGAGGGTCACCTCCCTCC  
ATGATACACACTGCACCTCTGGCCGTTGGATCCATCTCCAGGAAGCCCCACGACTGCC  
GCATCCAGGCCTTTCTCTTGGCATCTGTTCTGGAGGTTTCATCTCCATCTGCTATGAG  
AACATCCGCTCCCTCCAGGTCCAGATGTTGCCCTTACTAAGCGATGGTTTCACCGTCTC  
TTACCTACCATTCTGTCTCCAGACACTGACCCATGTGGGTCTCTTTTCTATTGTAC

40355 CTCCCCAGCTCTCCTCTCCCTCTTCTCCATCCACCCTACACTTGCTGCCAGACACAGATA  
GACCTTCTGGAAATAAAGTGTGCCCCATCAAGGCTGCTTGAAATCCTTGCTGATCCCTAC  
TGCCCATTTGACAGAGTCTGGAGGGAGGGTCACCTCCCTCCATGATACACACTGCACCTC  
TGGCCGGTGGATCCATCTCCAGGAAGCCCCACGACTGCCCGCATCCAGGCCTTTCTTTT  
TGCCATCTGTTCTGGAGGTTTCATCTCCATCTGCTATGAGAATCCGCTCCCTCCAG  
[G, C]  
TCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCGTCTCTTACCTACCATTCTGTCTC  
CAGACACTGACCCATGTGGGTCTCCTTTTCTATTGTACCTCTCATGAGACACCGACCA  
GTCTCCTTTATGATGTGATTGTTTCTGCACATCTCAACTTCTCCTGGGCCACAAGAAAA  
GATGTCACATCTTAACCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACAGC  
ACGTGCCCAATTACATTCAGTGTGAGAGGAGTGGAGAGGGGCATAGGAAGGCAAGAACG

40486 CAGAGTCTGGAGGGAGGGTCACCTCCCTCCATGATACACACTGCACCTCTGGCCGGTGGA  
TCCATCTCCAGGAAGCCCCACGACTGCCCGCATCCAGGCCTTTCTTTTGCCATCTGTT  
CCTGGAGGTTTCATCTCCATCTGCTATGAGAATCCGCTCCCTCCAGGTCCAGATGTT  
GCCCTTACTAAGCGATGGTTTACCCGCTCTCTTACCTACCATTCTGTCTCCAGACACTGA  
CCCATGTGGGTCTCCTTTTCTATTGTACCTCTCATGAGACACCGACCCAGTCTCCTTTA  
[T, C]  
GATGTGATTGTTTCTGCACATCTCAACTTCTCCTGGGCCACAAGAAAAGATGTACATC  
TTAACCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACAGCACGTGCCCAAT  
TCACATTCAGTGTGAGAGGAGTGGAGAGGGGCATAGGAAGGCAAGAACGCACACGATCTG  
CCCACATGCCTCCCCTCCCGCCCTTCTGATTGGGGATCTTTTCATCTACTACAAAACCA  
GCTGTCTTCCATGCTGCCCTTCCCTGATTTCTGGGTAGTCTGGGATGGGAGAATGGGG

40512 CTCCATGATACACACTGCACCTCTGGCCGGTGGATCCATCTCCAGGAAGCCCCACGACT  
GCCCGCATCCAGGCCTTTCTTTTGCCATCTGTTCTGGAGGTTTCATCTCCATCTGCTA  
TGAGAACATCCGCTCCCTCCAGGTCCAGATGTTGCCCTTACTAAGCGATGGTTTCACCG  
TCTCTTACCTACCATTCTGTCTCCAGACACTGACCCATGTGGGTCTCCTTTTCTATTG  
TACCTCTCATGAGACACCGACCCAGTCTCCTTTATGATGTGATTGTTTCTGCACATCTCA  
[A, C]  
CTTCTCTCTGGGCCACAAGAAAAGATGTACATCTTAACCTCCAGTCTCATCACAGCTT  
CCAGCAAGGGGGCTAAACACAGCACGTGCCCAATTACATTCAGTGTGAGAGGAGTGGAG  
AGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCACATGCCTCCCCTCCCGCCCTT  
CTGATTTGGGGATCTTTTCATCTACTACAAAACAGCTGTCCTTCCATGCTGCCCTTCCCT  
GATTTCTGGGTAGTCTGGGATGGGAGAATGGGGACAGTTGTGACCACGAGGAAGCAGAG

FIGURE 3, page 30 of 33

40622 CCATCTGCTATGAGAACATCCGCCTCCCTCCAGGTCCAGATGTTGCCTTTACTAAGCGAT  
GGTTTCACCGTCTCTTACCTACCATTCCTGTCTCCAGACACTGACCCATGTGGGTCTCCT  
TTTCTATTGTACCTCTCATGAGACACCGACCCAGTCTCCTTTATGATGTGATTGTTTCT  
GCACATCTCAACTTCTCCTGGGCCACAAGAAAAGATGTCACATCTTAACCCCTCCAGTCT  
CATCACAGCTTCCAGCAAGGGGGCTAAACACAGCACGTGCCAATTACATTCACTGAGA  
[- , A, G]  
GAGAGTGGAGAGGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCACATGCCTCCCCT  
CCCGGCCCTTCTGATTTGGGGATCTTTCATCTACTACAAAACCAGCTGTCTTCCATGCT  
GCCCTTCCCTGATTTCTGGGTAGTCTGGGATGGGAGAATGGGACAGTTGTGACCACGA  
GGAAGCAGAGGTGGGAGTTCTACAGGCCCCACAGGGCTCTCTGCCATTGGTCACTATCA  
GTTCCCAATCTTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAG

40654 GGTCCAGATGTTGCCTTTACTAAGCGATGGTTTCACCGTCTCTTACCTACCATTCCTGTC  
TCCAGACACTGACCCATGTGGGTCTCCTTTTCTATTGTACCTCTCATGAGACACCGACC  
CAGTCTCCTTTATGATGTGATTGTTTCTGCACATCTCAACTTCTCCTGGGCCACAAGAA  
AAGATGTCACATCTTAACCCCTCCAGTCTCATCACAGCTTCCAGCAAGGGGGCTAAACACA  
GCACGTGCCAATTACATTCACTGAGAGGAGAGTGGAGAGGGGCATAGGAAGGCAAGAA  
[T, C]  
GCACACGATCTGCCACATGCCTCCCCTCCCGGCCCTTCTGATTTGGGGATCTTTCATCT  
ACTACAAAACCAGCTGTCTTCCATGCTGCCCTTCCCTGATTTCTGGGTAGTCTGGGAT  
GGGAGAATGGGACAGTTGTGACCACGAGGAAGCAGAGGTGGGAGTTCTACAGGCCCCAC  
AGGGCTCTCTGCCATTGGTCACTATCAGTTCCTCAATCTTTCAAAATCAGGTTTGATGGC  
CAAGGAAACGCTGGTGAGAAACCAAAAGAGTTCTAGCTGGGTGTTGACCTCTTAGAG

40933 AGGGGCATAGGAAGGCAAGAACGCACACGATCTGCCACATGCCTCCCCTCCCGGCCCTT  
CTGATTTGGGGATCTTTCATCTACTACAAAACCAGCTGTCTTCCATGCTGCCCTTCCCT  
GATTTCTGGGTAGTCTGGGATGGGAGAATGGGACAGTTGTGACCACGAGGAAGCAGAG  
GTGGGAGTTCTACAGGCCCCACAGGGCTCTCTGCCATTGGTCACTATCAGTTCCTCAATC  
TTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAGAGTTCTAG  
[C, G]  
TGGGTGTTGACCTCTTTAGAGGCCCATCCCGCTAAAGAGGGTTTGGGCACAGCCTAAATG  
AGGGAGCTTTACAAAAGGGAAGCTCTGTGAAAACGTCAGGGTTATCGCAGCATCTCAGG  
AATGGGGACTAGGCAAGTCTTGGCTTGGTGATGGATGGTTACGGAGATCCTTTCCACTG  
ACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACCTCTACACCCTGCGGAACAGCCGATTG  
CTCCCAGACCAGGCTGAAGAAGCCCCCGAGGTGCTCTACTGGGCTGTCCAGTGGCAGCG

41171 TCTTTCAAAATCAGGTTTGATGGCCAAGGAAACGCTGGTGAGAAACCAAAAGAGTTCT  
AGCTGGGTGTTGACCTCTTTAGAGGCCCATCCCGCTAAAGAGGGTTTGGGCACAGCCTAA  
ATGAGGGAGCTTTACAAAAGGGAAGCTCTGTGAAAACGTCAGGGTTATCGCAGCATCTC  
AGGAATGGGGACTAGGCAAGTCTTGGCTTGGTGATGGATGGTTACGGAGATCCTTTCCA  
CTGACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACCTCTACACCCTGCGGAACAGCCGA  
[T, C]  
TGCTCCAGACCAGGCTGAAGAAGCCCCCGAGGTGCTCTACTGGGCTGTCCAGTGGCAG  
CGGGAGTCGGCCTTGTGGTGGGCATCATCGGGACAGTGGTCTGGAGGTGCAGGAGAGGTG  
GCCGGAGGGAAGATCCTCAATGAGCCTGCGCACTGTGCCCTCTAGGCCCCGGGGTGGG  
TCCTCACCTTAACTCCCTATAGCCACTCTCTTACCGCCCTCTGCCCCAGCCACTCCCG  
GCCACCAGGACATGCTTCAATGCCTGGTGCCATAGGAAGCCAGGGGACAGTCACAACTT

41379 TGGTGATGGATGGTTACGGAGATCCTTTCCACTGACCCCCGCTCCTCCTCCACAGGAGA  
GCCCTTACCTCTACACCCTGCGGAACAGCCGATTGCTCCAGACCAGGCTGAAGAAGCCC  
CCGAGGTGCTCTACTGGGCTGTCCAGTGGCAGCGGGAGTCGGCCTTGTGGTGGGCATCA  
TCGGGACAGTGGTCTGGAGGTGCAGGAGAGGTGGCCGGAGGGAAGATCCTCCAATGAGCC  
TGCGCACTGTGGCCTCTAGGCCCGGGGGTGGGTCTCACCTTAACTCCCTATAGCCAC  
[T, C]  
CTCTTACCGCCCTCTGCCCCAGCCACTCCCGGCCACCAGGACATGCTTCAATGCCTGGT  
GCCATAGGAAGCCAGGGGACAGTCACAACTTCTTGGGGCTGGGCTTCTTCCAGGCCTA  
TGCTCCTGGAATGGATACATTTAAATAAAGTCCAAAGCTATTTTATTCTGGGTTTGCTT  
GCGTGAAGCACTCACCTTCCATCTCTTGTGCAGCCAGGTGTGGGAGCTGCCACTTTTGT  
TGGCCTGCCTCCAGCAGGGCTGCCAAGCCACGACCAACCAGAGCCCAACTGCCTGCCA

41388 ATGGTTACAGGAGATCCTTTCCACTGACCCCCGCTCCTCCTCCACAGGAGAGCCCTTACC  
TCTACACCCTGCGGAACAGCCGATTGCTCCAGACCAGGCTGAAGAAGCCCCCGAGGTGC  
TCTACTGGGCTGTCCAGTGGCAGCGGGAGTCGGCCTTGTGGTGGGCATCATCGGGACAG

FIGURE 3, page 31 of 33



TGGTCTGGAGGTGCAGGAGAGGTGGCCGGAGGGAAGATCCTCCAATGAGCCTGCGCACTG  
TGGCCCTCTAGGCCCGGGGTGGGTCTCACCCTAAACTCCCTATAGCCACTCTCTTAC  
[C, T]  
GCCCTCTGCCCCAGCCACTCCCGGCCACCAGGACATGCTTCAATGCCTGGTGCCATAGGA  
AGCCAGGGGACAGTCACAACCTTCTTGGGGCCTGGGCTTCTCCAGGCCTATGCTCCTGG  
AATGGATACATTTAAATAAAGTCCAAAGCTATTTTATTCCTGGGTTTGCCTGCGTGAAGC  
ACTCACCTTCCATCTCTTGTGCAGCCAGGTGTGGGAGCTGCCACTTTTTGTGGCCTGCC  
TCCAGCAGGGCTGCCCAAGCCACGACCAACCAGAGCCAAACTGCCTGCCACCACGAGCA

41880 ATCTCTTGTGCAGCCAGGTGTGGGAGCTGCCACTTTTTGTGGCCTGCCTCCAGCAGGGC  
TGCCCAAGCCACGACCAACCAGAGCCAAACTGCCTGCCACCACGAGCATATCCTCAAGT  
CACCAAACCCACTATTTCAAAGGCAGAAAAAATGCTGGTCAACAGGTGGTGGCTGGAATT  
TTGGAGCTGGCTGGTTGCCATTCACTCCAATCCAACACATACCTATTAAGCAACTGTTTT  
GTATCCAGGACAATGCGAAGCACTGAGGTGCCTCCTAGGCTGTGCATGTGCGAGCCTGGC  
[A, C]  
GAGAGGTCAAACCTCCTTCAATAACCAAGAAGCCACGTGATGATGTGTAACACTAGGGCA  
TCAGTAGGTAAATGTGTCTGATTGTTTTAAAGAATAGAAAGGGTCTTCGGGGAAAGTTT  
CTTGGGGGAGAGCAACCTTCACATGTCTATTTGGGAAAAGGAATAAAAAATGATTGGGAC  
ACAAATACCTCCTATATTCTCAACCTGATTTCTCAAGGTGCTAAATTTAGGAAAAAATT  
CCTATTTCTATATGCCAGGTTTCTGAGGGAAAAGTCTGAGAGAGTCTGAAAATATGGGCT

42278 AAAGGGTCTTCGGGGAAAGTTTCTTGGGGGAGAGCAACCTTCACATGTCTATTTTGGGAA  
AAGGAATAAAAAATGATTGGGACACAAATACCTCCTATATTCTCAACCTGATTTTCTCAA  
GGTGCTAAATTTAGGAAAAAATTCCTATTTCTATATGCCAGGTTTCTGAGGGAAAACTA  
GAGAGAGTCTGAAAATATGGGCTGCATTCAGTCAAGCCCTGCTAGGGGCGAGGCCCCGTG  
CTGGAGGCCTTCACAGATGGTCTCTTTTATGCTGCACAAAAGCCAGGGAGGGGTAAA  
[G, A]  
GGAAAATCTTTGAAAATAGAAGTGATGCTTGCACAAACCCGTGAATGTACTAAACGCCGC  
GAATTGTTCCATTTAAATGATTAATTGTGTATCATGTGAATTTCACTTCAATAAAAAAG  
AATCCAGGGAGGTAGACATCATCTGCATTGTAAACCTCTCTGATCCTGAAGTCCGGGA  
TGATAAAGAGCCTGAGTCACAATCCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTG  
CCTTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGTAAAACTGGAGTAAGAGT

42339 AGGAATAAAAAATGATTGGGACACAAATACCTCCTATATTCTCAACCTGATTTTCTCAAG  
GTGCTAAATTTAGGAAAAAATTCCTATTTCTATATGCCAGGTTTCTGAGGGAAAACTAG  
AGAGAGTCTGAAAATATGGGCTGCATTCAGTCAAGCCCTGCTAGGGGCGAGGCCCCGTG  
TGGAGGCCTTCACAGATGGTCTCTTTTATGCTGCACAAAAGCCAGGGAGGGGTAAAA  
GGAAAATCTTTGAAAATAGAAGTGATGCTTGCACAAACCCGTGAATGTACTAAACGCCGC  
[G, A]  
AATTGTTCCATTTAAATGATTAATTGTGTATCATGTGAATTTCACTTCAATAAAAAAGA  
ATCCAGGGAGGTAGACATCATCTGCATTGTAAACCTCTCTGATCCTGAAGTCCGGGAT  
GATAAAGAGCCTGAGTCACAATCCCGGATGCAACACTGAAATGCTGTGCCCTGAAGCTGC  
CTTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGTAAAACTGGAGTAAGAGTA  
CACATTTTGTCTTATCTACGGCGCTGCTGAAAAATAAGGAACCGTGTGTGAACCTCTAAC

42612 CAACACCGTGAATGTACTAAACGCCGCAATTGTTCCATTTAAATGATTAATTGTGTAT  
CATGTGAATTTCACTTCAATAAAAAAGAATCCAGGGAGGTAGACATCATCTGCATTGTAA  
ACCTCTCTCTGATCCTGAAGTCCGGGATGATAAAGAGCCTGAGTCACAATCCCGGATGCA  
ACACTGAAATGCTGTGCCCTGAAGCTGCCCTTCGCCAGCCTGAGCCAGTGTCCAGGCTC  
TGCATCTGTAAAACTGGAGTAAGAGTACACATTTTGTCTATCTACGGCGCTGCTGAAA  
[A, G]  
ATAAGGAACCGTGTGTGAACCTCTAACTCTAAAATGCTGCACAACTGAAAATGGCCTTTT  
TCCTCGGTGAAGAGTTGGGATAAGGCCAGACTGTTGGGGAAGATGTGAGACCCAGAGAT  
GAGTTTGGGGAATGGGGTAATAACATATGGGTGGAGAGTGCCCGCCTTCTCTCAGGGA  
GGTTCATCACCTTATCTCTTTCTGTCAACAGAGAACCCGGAGGACCTATACCCAGTTC  
CGTGTCTCTTGGGCTTCAGTGTCTGTTTCTATACAATGGGAACAGCATGCATTCCCCTG

42817 TGCCTTCGCCAGCCTGAGCCAGTGTCCAGGCTCTGCATCTGTAAAACTGGAGTAAGA  
GTACACATTTTGTCTATCTACGGCGCTGCTGAAAAATAAGGAACCGTGTGTGAACCTCT  
AACTCTAAATGCTGCACAACTGAAAATGGCCTTTTCTCCTCGGTGAAGAGTTGGGATAAG  
GCCCAGACTGTTGGGGAAGATGTGAGACCCAGAGATGAGTTTGGGGAATGGGGTAATAA  
CATATGGGTGGAGAGTCCCCGCTTCTCTCAGGGAGGTTTCATCACCTTATCTCTTTCTG  
[T, G]

FIGURE 3, page 32 of 33

[illegible]